

# Theory and Application of a Pure-sampling Quantum Monte Carlo Algorithm

by

**Egor Ospadov**

M. Sc. (Theoretical Physics) Department of Physics, Brock University, Canada

B. Sc. (Honours Physics) Department of Physics and Astronomy,

University of Waterloo, Canada

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

in

The Faculty of Mathematics and Sciences

Department of Physics

BROCK UNIVERSITY

April 16, 2016

2016 © Egor Ospadov

# Abstract

The objective of pure-sampling quantum Monte Carlo is to calculate physical properties that are independent of the importance sampling function being employed in the calculation, save for the mismatch of its nodal hypersurface with that of the exact wave function. To achieve this objective, we describe a pure-sampling algorithm that combines features of forward-walking methods of pure-sampling and reptation quantum Monte Carlo. The importance sampling is performed by using a single-determinant basis set composed of Slater-type orbitals. We implement our algorithm by systematically increasing an algorithmic parameter until the properties sampled from the electron distributions converge to statistically equivalent values, extrapolated in the limit of zero time-step. In doing so, we are able to unambiguously determine the values for the ground-state fixed-node energies and one-electron properties of various molecules. These quantities are free from importance sampling bias, population control bias, time-step bias, extrapolation-model bias, and the finite-field approximation. We applied our algorithm to the ground-states of lithium hydride, water and ethylene molecules, and found excellent agreement with the accepted literature values for the energy and a variety of other properties for those systems. Some of our one-electron properties of ethylene had not been calculated before at any level of theory. In a detailed comparison, we found reptation quantum Monte Carlo, our closest competitor, to be less efficient by at least a factor of two. It requires different sets of time-steps to accurately determine the ground-state energy and one-electron properties, whereas our algorithm can achieve the same objective by using a single set of time-step values.

# Acknowledgements

I take this opportunity to express my utmost gratitude to my supervisor, Dr. S. M. Rothstein, for mentoring me through all these years. He shared his immense knowledge with me in a professional and enjoyable manner. He made this an unforgettable experience that I will always remember. Without him, I would not be where I am right now or in the future.

I want to show my appreciation to the professors and the staff in the Physics Department, for helping me grow, both academically and professionally. I want to express my appreciation to my friends, who kept my spirits up and wits sharp. Finally, I am very grateful to my family for their ongoing support during all these years. They provided me with unconditional support, inspiration, and courage whenever I needed them.

Thank you.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theory</b>	<b>6</b>
2.1	Modified Schrödinger Equation in Imaginary Time . . . . .	6
2.2	Evolution Through Imaginary Time . . . . .	8
2.3	A Pure-sampling Quantum Monte Carlo Method . . . . .	9
2.4	Fixed-node Approximation . . . . .	15
2.5	Estimating the Energy . . . . .	16
2.6	Truncations . . . . .	17
<b>3</b>	<b>Electronic Properties</b>	<b>18</b>
3.1	Electric Moments . . . . .	18
3.2	Diamagnetic Shielding and Susceptibility . . . . .	20
3.3	Electric Fields and Electric Field Gradients . . . . .	21
3.4	Polarizabilities . . . . .	22
<b>4</b>	<b>Results and Discussion</b>	<b>24</b>
4.1	Application to Lithium Hydride . . . . .	25
4.2	Application to Water . . . . .	31
4.3	Application to Ethylene . . . . .	36
4.4	Comparison with RQMC . . . . .	40
<b>5</b>	<b>Conclusion</b>	<b>44</b>
<b>6</b>	<b>References</b>	<b>46</b>
<b>A</b>	<b>Appendix</b>	<b>54</b>
A.1	Derivation of the Modified Schrödinger Equation . . . . .	54
A.2	Derivation of the Energy Estimator . . . . .	57

A.3	Graphs of the Ground-state Energy and Electronic Properties of Lithium	
	Hydride . . . . .	58
A.4	Graphs of the Ground-state Energy and Electronic Properties of Water	72
A.5	Graphs of the Ground-state Energy and Electronic Properties of Ethylene	90

## List of Figures

1	Visual representation of the process used to generate $X$ and $Y$ . . . .	11
2	Schematic representation of an electron distribution . . . . .	13
3	Illustration of geometries for lithium hydride, water and ethylene . . .	19
4	Ground-state energy of lithium hydride . . . . .	26
5	Dipole moment of lithium hydride . . . . .	27
6	Breakdown of the dipole moment of lithium hydride . . . . .	28
7	Ground-state energy of water . . . . .	31
8	Dipole moment of water . . . . .	32
9	Ground-state energy of ethylene . . . . .	36
10	Quadrupole moment of ethylene . . . . .	37
11	Properties of lithium hydride calculated using RQMC . . . . .	40
12	Properties of lithium hydride calculated using PSQMC . . . . .	41
13	CPU time versus time-step for RQMC and PSQMC . . . . .	42
14	Metropolis acceptance rates versus time-steps for RQMC and PSQMC	43

## List of Tables

1	Ground-state energy and dipole moment of lithium hydride . . . . .	25
2	Ground-state energy and electronic properties of lithium hydride . . .	30
3	Ground-state energy and dipole moment of water . . . . .	33
4	Additional literature results for water polarizabilities . . . . .	34
5	Ground-state energy and electronic properties of water . . . . .	35
6	Ground-state energy and quadrupole moment of ethylene . . . . .	38
7	Ground-state energy and electronic properties of ethylene . . . . .	39

## List of Abbreviations

<b>ADF</b>	Amsterdam Density Functional
<b>CC</b>	Coupled Cluster
<b>CCSD</b>	Coupled Cluster with single and double excitation
<b>CI</b>	Configuration Interaction
<b>CPU</b>	central processing unit
<b>DFT</b>	Density Functional Theory
<b>DMC</b>	diffusion Monte Carlo
<b>FCI</b>	full Configuration Interaction
<b>GPU</b>	graphics processing unit
<b>HF</b>	Hartree-Fock
<b>PSQMC</b>	pure-sampling quantum Monte Carlo
<b>QMC</b>	quantum Monte Carlo
<b>QZ4P</b>	quadruple zeta with 4 polarization functions
<b>RQMC</b>	reptation quantum Monte Carlo
<b>RQMC-MH</b>	Metropolis-Hastings version of reptation quantum Monte Carlo
<b>SCF</b>	Self-Consistent Field
<b>STO</b>	Slater-type orbital



## List of Nomenclature

**basis set** a linear combination of functions that describe molecular orbitals

**correlation energy** difference between the exact energy and the Hartree-Fock limit

**exchange node** the wave function equals to zero when two electrons with the same spin coalesce

**extrapolation-model bias** bias due to an ambiguous choice of the fitting model

**finite-field approximation** applying a small perturbation in the form of an external electric or magnetic field to calculate quantities in the limit of zero field

**fixed-node approximation** the nodes of the importance sampling function are imposed upon the unknown, exact ground-state wave function

**importance sampling bias** bias introduced by the importance sampling function to properties calculated from the pure distribution

**importance sampling function** guides electrons away from unimportant regions of space and towards important ones

**local energy** the energy of a particular electron configuration

**mixed distribution** product of the importance sampling function and the unknown, exact ground-state wave function

**nodal hypersurface** a multi-dimensional surface that contains exchange nodes

**population control bias** bias arising from artificially controlling the amount of walkers in diffusion Monte Carlo calculations

**pure distribution** square of the unknown, exact ground-state wave function

**serial correlation** similarity between observations over an interval of time

**statistical weight** relative weight of a particular sample or configuration

**time-step bias** failure of observations to be constant with time-step, which can be ameliorated with extrapolation

**variational-bound** the energy must be equal to or larger than the exact ground-state energy

# 1 Introduction

The conception of quantum mechanics in the early 20<sup>th</sup> century paved the way for a new era in both experimental and theoretical physics. The ability of quantum mechanics to explain previously-unsolved phenomena by what is now referred to as classical physics and predict new observations was hailed as revolutionary. One such famous problem that could not be explained by classical physics was the photoelectric effect, observed by Heinrich Hertz in 1887 and explained by Albert Einstein in 1905 using quantum mechanics.<sup>1,2</sup> Additionally, the Compton scattering experiment conducted by Arthur Compton in 1923 could only be explained by quantum mechanics.<sup>3</sup> The solutions to these problems produced the notion of wave-particle duality: light can be treated as a collection of particles, called photons, and particles can exhibit both wave-like and particle-like properties. Today, quantum mechanics is applied to explain and predict phenomena not just in the field of physics but also in the fields of chemistry, biology, and material science.

While quantum mechanics is a remarkable theory, it is not without its shortcomings. There are relatively few systems that can be solved exactly using quantum mechanics. The problem arises in the complexity of the mathematics required to solve the Schrödinger equation, the mathematical backbone of quantum mechanics, when it is applied to atoms, molecules, soft-matter and solid-state systems. As the size of the system (e.g., number of electrons) grows, the equations associated with the system become more difficult to solve analytically. To solve larger systems, initial attempts involved using approximations, such as the variational method or perturbation theory, to reduce the complexity of the associated equations. While these methods provided us with a modest increase in the size of the system that could be treated using quantum mechanics, they did so at the expense of the accuracy of the results. Ultimately, these methods led to a deeper understanding of quantum mechanics, but the goal of being able to solve any possible system analytically eludes us to this day.

After coming to terms with our current inability to solve the Schrödinger equation exactly, the next natural step was to develop approximation methods capable of giving us a solution with desired accuracy. Surging technological advances in the field of high-performance computing have greatly aided this approach by providing the necessary computing power to implement these methods. One such widely-used and successful method is called quantum Monte Carlo (QMC), which is an extension of the classical Monte Carlo scheme applied to solve the Schrödinger equation.<sup>4-6</sup>

QMC is a set of computational methods used to study quantum systems that can be described with a wave function, by obtaining the solution to the associated Schrödinger equation. Fundamentally, it can be applied to study the ground<sup>7-9</sup> or excited<sup>8-11</sup> states of any bosonic or fermionic system. Our discussion will focus on methods that are used to study the ground-state of atoms and molecules. In many cases, QMC methods not only solve the Schrödinger equation to obtain the energy but also provide ways to estimate, or sample, various experimentally-important properties such as the dipole moment, polarizabilities and vibrational frequencies.

Depending on the type of QMC method, the solution to the Schrödinger equation can yield either a *mixed distribution* of electrons,  $\Psi\Phi_0$ , a *pure distribution* of electrons,  $\Phi_0^2$ , or both. Here,  $\Phi_0$  is the unknown, putatively exact ground-state wave function. These distributions can be generated using an *importance sampling* function,  $\Psi$ , which guides electrons away from unimportant regions of space and towards important ones. There are methods that can sample directly from  $\Phi_0$ , bypassing the requirement for  $\Psi$ , but their range of application is currently limited.<sup>7</sup> Except for node-less systems, such as the hydrogen atom or the ground-state of helium atom, these distributions are generated with a minimal bias originating from the mismatch between the nodal hypersurface of the importance sampling function and that of the exact ground-state wave function. The type of available distribution is important, because it dictates what properties can be obtained from it using importance sampling

without underlying bias, except due to nodal errors.

Diffusion Monte Carlo (DMC) with importance sampling is the most widely-used quantum Monte Carlo method.<sup>6,7</sup> It is capable of sampling only from the mixed distribution of electrons. Using DMC, only the fixed-node energy of the system can be determined directly, albeit with a fixed-node bias and a time-step bias, which is readily dealt with. Any property represented by an operator that does not commute with the Hamiltonian is biased more severely. This is a direct result of DMC's inability to generate and sample from the pure distribution of electrons. The mixed estimate of such a property  $\hat{A}$  introduced in DMC,  $\langle \Psi | \hat{A} | \Phi_0 \rangle / \langle \Psi | \Phi_0 \rangle$ , yields degraded results due to its inherent and large importance sampling function bias.<sup>12</sup> The actual desired result is given by  $\langle \Phi_0 | \hat{A} | \Phi_0 \rangle / \langle \Phi_0 | \Phi_0 \rangle$ .

Reptation quantum Monte Carlo (RQMC), originally developed by S. Baroni and S. Moroni<sup>13</sup>, is a straight-forward and computationally-efficient method used to sample both mixed and pure distributions of electrons. Hence, RQMC can be used to sample both the ground-state energy and ground-state properties associated with operators that commute with the position operator.<sup>13</sup> Furthermore, RQMC does not suffer from importance sampling bias or walker population control bias, both of which occur in DMC. Disadvantages of RQMC are a presence of a large, albeit controllable time-step bias and a small, but uncontrollable bias stemming from sampling the inexact nodal hypersurface of the importance sampling function. These biases are not inherent to RQMC alone, and in fact appear in DMC<sup>14</sup> as well as its derivatives. Additionally, a Metropolis-Hastings version of RQMC (RQMC-MH) was proposed by Yuen *et al.*<sup>15</sup> to ameliorate the failure of microscopic reversibility in the limit of zero time-step and ensure the convergence to the intended distribution.

There are other methods beside quantum Monte Carlo that are designed to give an approximate solution to the Schrödinger equation for atoms and molecules. Some of these widely-used methods include Density Functional Theory (DFT)<sup>16–19</sup>, Cou-

pled Cluster (CC)<sup>20–24</sup>, and Configuration Interaction (CI)<sup>25–29</sup>. QMC methods have intrinsic disadvantages when compared to these methods, such as statistical error in the calculations and various biases discussed previously. QMC methods also present a big advantage over these methods. QMC is variationally-bound, requires insignificant computational resources and has favourable scaling, but it does require a large investment of computing time. For the sake of comparison, QMC scales as  $O(s^3)$ , DFT scales as  $O(s^3)$ <sup>18</sup>, CC scales as  $O(s^7)$ <sup>24</sup> and CI, which is variationally-bound, scales as  $O(s^{10})$ <sup>28</sup>. Here,  $s$  is the number of electrons in the atom or a molecule. Setting the issue of accuracy aside, the scaling establishes limitations to the practical applications of these methods. QMC methods have the potential to be applied to very large systems, such as molecular clusters, polymers and proteins, due to their favourable scaling being on par with DFT. On the other hand, CC and CI methods are capable of studying small molecules but they become too computationally infeasible when applied to large systems, demanding exhaustive CPU and IO resources.

The algorithm presented in this work rests on a single, uncontrollable approximation: the fixed-node approximation, which plagues many other quantum Monte Carlo methods.<sup>6,14,30–32</sup> Recently, Booth *et al.*<sup>33</sup> introduced a new QMC method that takes elements from diffusion Monte Carlo and full Configuration Interaction (FCI)<sup>28</sup> and does not rest upon the fixed-node approximation. In this method, they create walkers with a positive or negative sign and evolve them in the Slater-determinant space according to a set of steps: spawning (creating a walker), death (destroying a walker) and annihilation (destroying a pair of walkers with opposite signs that are found on the same determinant). They show that their method converges to the FCI solution by using two independent estimators of the correlation energy and applying it to multiple molecules. The only source of error in their calculation arises from the incompleteness of the basis set, which can be ameliorated by using larger basis sets and/or extrapolating towards the complete basis set limit.<sup>33</sup> Their favourable memory

requirements allow them to minimize this error and explore larger molecules. They also show that their method can calculate excited states without a loss of efficiency.<sup>34</sup> However, in common with FCI, their scope is still limited due to their dependence on multiple, large basis sets and factorial scaling, albeit with a smaller pre-factor.<sup>33</sup>

This work focuses on a novel pure-sampling quantum Monte Carlo (PSQMC)<sup>35</sup> method that is capable of sampling both mixed and pure distributions simultaneously in order to estimate properties that are free from population control bias, importance sampling bias, time-step bias, extrapolation-model bias and the finite-field approximation. The results obtained using this method are of very high quality, both in terms of accuracy and precision. The systematic approach that we take in this work and the observed convergence of the results provides us with confidence in our choice of algorithmic parameters. We show this by applying our new method to lithium hydride, water and ethylene molecules to calculate their ground-state energies, electronic properties and polarizabilities. Furthermore, our results show that this method is computationally superior to RQMC, its nearest competitor.

In the following section, we describe the theory behind our new pure-sampling algorithm. Afterwards, we discuss the quantities that we calculated in this work and present their respective estimators. Subsequently, we give an account of the method that we used to establish the convergence of our algorithmic parameters. Next, we present and discuss the results of our application and make a comparison with RQMC. Finally, we conclude this work with some final remarks. The Appendices include key derivations presented throughout this work and all relevant figures that were used to calculate the results from our application.

## 2 Theory

In this section we describe the theory behind our pure-sampling quantum Monte Carlo (PSQMC) algorithm.<sup>35</sup> The objective of our algorithm is to calculate various physical properties of different molecules free of as much bias as possible. We accomplish this by successfully simulating a solution to a modified Schrödinger equation, generating both mixed and pure distributions of electrons simultaneously and then sampling them to obtain the desired properties. This is achieved by using drift and diffusion moves, mathematical moves performed in imaginary time that do not carry any physical significance, to generate main and sub paths consisting of electron configurations. After evoking a simple Metropolis decision, the sub paths allow sampling of properties from the mixed distribution (at their ends) and pure distribution (at their middle). The main path reduces the serial correlation of the electron configurations where the pure distribution is sampled by providing an imaginary time separation between them.

### 2.1 Modified Schrödinger Equation in Imaginary Time

As a first step towards achieving our objective, we introduce a modified version of the Schrödinger equation expressed in imaginary time ( $t \rightarrow it$ ) and atomic units:

$$-\frac{1}{2}\nabla^2 f(\mathbf{r}) + \nabla \cdot (f(\mathbf{r})\mathbf{F}(\mathbf{r})) + (E_{loc}(\mathbf{r}) - E_{trial}) f(\mathbf{r}) = -\frac{\partial f(\mathbf{r})}{\partial t} = 0 \quad (1)$$

A solution to this equation is the *mixed distribution*:

$$f(\mathbf{r}) = \Psi(\mathbf{r})\Phi_0(\mathbf{r}) \quad (2)$$

that is only obtainable in the limit of infinite imaginary time. A detailed derivation of the modified Schrödinger equation and its solution can be found in Appendix 1.

In the equations written above,  $\mathbf{r}$  is the set of  $3n$  coordinates of  $n$  electrons.  $\Psi(\mathbf{r})$

is the *importance sampling* function, an approximate wave function used to guide electrons away from unimportant regions of space and towards important regions, consequently reducing the variance of simulated quantities. In a single-reference approach for closed-shell systems, such as the ones studied in this work,  $\Psi(\mathbf{r})$  is a product of two Slater determinants, one of which contains all spin-up electrons and the other contains all spin-down electrons.  $\Phi_0(\mathbf{r})$  is the unknown exact ground-state wave function, albeit biased by the fixed-node approximation.

$\mathbf{F}(\mathbf{r})$  is the *drift velocity* (also known as *quantum force*) and it is defined as:

$$\mathbf{F}(\mathbf{r}) = \frac{\nabla \Psi(\mathbf{r})}{\Psi(\mathbf{r})} \quad (3)$$

$E_{loc}(x)$  is the *local energy*. It represents the energy of an electron configuration at a specific instance of imaginary time. It is given by:

$$E_{loc}(\mathbf{r}) = \frac{\hat{H}\Psi(\mathbf{r})}{\Psi(\mathbf{r})} \quad (4)$$

where  $\hat{H}$  is the molecular Hamiltonian within the Born-Oppenheimer approximation<sup>36</sup>, expressed in atomic units as:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^n \nabla_i^2 - \sum_{i=1}^N \sum_{j=1}^n \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_j|} + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^N \sum_{j>i}^N \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \quad (5)$$

Here,  $n$  and  $N$  are the number of electrons and nuclei in the molecule, respectively.  $Z$  is the charge of a nucleus.  $\mathbf{r}$  represents the position of an electron and  $\mathbf{R}$  represents the position of a nucleus.

Lastly,  $E_{trial}$  is the trial energy that corresponds to the exact ground-state energy of the system. Our methodology does not depend on the value of  $E_{trial}$ , because the factors that contain it cancel out exactly with each other (see Equation (20)).



## 2.2 Evolution Through Imaginary Time

Recall that our solution to the Schrödinger equation is only valid in the limit of infinite imaginary time. More formally, we can write  $f$  at any imaginary time  $t$  as:

$$f(\mathbf{r}, t) = \Psi(\mathbf{r})\Phi(\mathbf{r}, t) \quad (6)$$

In order to obtain our desired mixed distribution,  $\Psi(\mathbf{r})\Phi_0(\mathbf{r})$ , we must find a way to evolve  $\Phi(\mathbf{r}, t = 0)$  in imaginary time to achieve  $\Phi(\mathbf{r}, t \rightarrow \infty) = \Phi_0(\mathbf{r})$ . To accomplish this, we can express the time evolution of  $f(\mathbf{r}, t)$  in integral form as:

$$f(\mathbf{r}', t + \tau) = \int G(\mathbf{r} \rightarrow \mathbf{r}'; \tau) f(\mathbf{r}, t) d\mathbf{r} \quad (7)$$

Here,  $\tau$  is the *time-step*, a small fraction of imaginary time.  $G(\mathbf{r} \rightarrow \mathbf{r}'; \tau)$  is the Green's function (or *propagator*) that describes the transition probability of moving a set of electron configurations from  $\mathbf{r}$  to  $\mathbf{r}'$ . It obeys the boundary condition  $G(\mathbf{r} \rightarrow \mathbf{r}'; 0) = \delta(\mathbf{r}' - \mathbf{r})$ . The problem is that an exact solution to this Green's function is not known. A remedy to this problem is to employ an approximate Green's function.

We start by assuming that  $\tau$  is small, the so-called *short-time* approximation, and that for such  $\tau$  the drift velocity is constant,  $\mathbf{F}(\mathbf{r}') \approx \mathbf{F}(\mathbf{r})$ . We can derive an approximate short-time Green's function by applying the Suzuki-Trotter expansion<sup>37,38</sup> to obtain the following equation:<sup>6,39</sup>

$$G(\mathbf{r} \rightarrow \mathbf{r}'; \tau) \approx (2\pi\tau)^{-\frac{3}{2}n} e^{-\tau \left( \frac{E_{loc}(\mathbf{r}') + E_{loc}(\mathbf{r})}{2} - E_{trial} \right)} e^{-\frac{(\mathbf{r}' - \mathbf{r} - \tau \mathbf{F}(\mathbf{r}))^2}{2\tau}} \quad (8)$$

This equation is a solution to Equation (1), accurate to second order in  $\tau$ , although the energy obtained using this equation is accurate to first order in  $\tau$ .<sup>40</sup> By evolving the electron configuration in imaginary time using this equation and a very small time-step, we will achieve our desired mixed distribution provided that the electron distribution advanced through enough imaginary time ( $t \rightarrow \infty$ ). Approximate short-time Green's functions of higher accuracy that reduce the time-step bias have been

reported by our group in the past.<sup>41,42</sup> However, the methodology reported in this work makes the task of finding more accurate short-time Green's functions obsolete.

The short-time Green's function is simulated by implementing *drift and diffusion* moves according to:

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \tau \mathbf{F}(\mathbf{r}_i) + \sqrt{\tau} \chi \quad (9)$$

where  $\tau \mathbf{F}(\mathbf{r}_i)$  is the *drift term*,  $\sqrt{\tau} \chi$  is the *diffusion term*, and  $\chi$  is a random number generated from a standard  $3n$ -dimensional normal distribution. This process generates variationally-distributed,  $\Psi^2$ , configurations of electrons using these mathematical moves that have no underlying physical significance.

### 2.3 A Pure-sampling Quantum Monte Carlo Method

Starting with  $x_0$ <sup>a</sup> and using drift and diffusion moves described in the previous section, we generate a path  $X$ , defined as:

$$X = \{x_{-L}, x_{-L+1}, \dots, x_{-1}, x_0, x_1, \dots, x_{L-1}, x_L\} \quad (10)$$

Here, each  $x_i$  is a set of  $3n$  electron coordinates and is a random sample from the variational distribution. The path is generated starting from  $x_0$  and doing two sets of  $L$  drift and diffusion moves. The total length of path  $X$  is  $2L + 1$  and the middle is defined at  $x_0$ . We wish to sample from the target distribution of this path, which we take to be:

$$\begin{aligned} \hat{\Pi}(X) &\propto G(x_{-L+1} \rightarrow x_{-L}; \tau) \dots G(x_{-1} \rightarrow x_{-2}; \tau) G(x_0 \rightarrow x_{-1}; \tau) \Psi^2(x_0) \\ &\times G(x_0 \rightarrow x_1; \tau) G(x_1 \rightarrow x_2; \tau) \dots G(x_{L-1} \rightarrow x_L; \tau) e^{-S(X)} \end{aligned} \quad (11)$$

---

<sup>a</sup>The initial electron configuration,  $x_0$ , used in the first iteration of our algorithm is generated pseudo-randomly. Each subsequent  $x_0$  comes from the path that is chosen by the Metropolis decision (see Equation (20)). Furthermore, the entire path  $X$  is never generated from scratch again, thereby increasing the computational efficiency of the algorithm by only requiring the new paths  $Z$  and  $Y$  to be generated at each subsequent iteration (see Figure 1).

where

$$G(x \rightarrow x'; \tau) \propto \exp \left[ \frac{-1}{2\tau} \left| x' - x - \tau \frac{\nabla \Psi(x)}{\Psi(x)} \right|^2 \right] \quad (12)$$

is the Green's function associated with the drift and diffusion moves<sup>b</sup>, and

$$S(X) = \tau \left[ \frac{1}{2} E_{loc}(x_{-L}) + \sum_{i=-L+1}^{L-1} E_{loc}(x_i) + E_{loc}(x_0) + \frac{1}{2} E_{loc}(x_L) \right] \quad (13)$$

is the sum of the local energies accumulated across the path. The individual local energies are calculated using Equation (4).

The formalism of path  $X$ , the target distribution, and the Green's function presented thus far are not unique to our method alone. They appear in other quantum Monte Carlo methods that are capable of pure sampling, such as reptation quantum Monte Carlo (RQMC)<sup>13,15</sup> and path integral ground state (PIGS)<sup>43–45</sup>. However, the rest of the formalism that follows separates us from these and other existing methods.

In order to sample from  $\hat{\Pi}(X)$ , we generate a new path  $Y$  from  $X$  using an intermediate path  $Z$ . The importance of this intermediate path will be discussed at the end of this section. This procedure of generating  $Y$  is as follows. First, using the middle of  $X$ ,  $x_0$ , as our starting point, do  $L$  drift and diffusion moves to generate  $Z$ , defined as:

$$Z = \{z_1, z_2, \dots, z_{L-1}, z_L\} \quad (14)$$

Next, starting at  $z_L$  and using a single drift and diffusion move, generate a new middle  $y_0$ . Finally, starting at the new middle,  $y_0$ , do two sets of  $L$  drift and diffusion moves to generate the rest of path  $Y$ , denoted as:

$$Y = \{y_{-L}, y_{-L+1}, \dots, y_{-1}, y_0, y_1, \dots, y_{L-1}, y_L\} \quad (15)$$

---

<sup>b</sup>Equation (12) is a simplified form of Equation (8), where all the constant terms are collected into the proportionality sign. Using Equation (12), instead of Equation (8), in the derivation presented in this section will yield an identical answer, because the constant terms will cancel out exactly upon substitution. This is also the reason as to why our methodology does not depend on the value of  $E_{trial}$  (see Equation (20)), precisely because it is one of these factors that cancel out exactly.

Similarly to  $X$ , our new path  $Y$  has the following target distribution:

$$\begin{aligned} \hat{\Pi}(Y) &\propto G(y_{-L+1} \rightarrow y_{-L}; \tau) \dots G(y_{-1} \rightarrow y_{-2}; \tau) G(y_0 \rightarrow y_{-1}; \tau) \Psi^2(y_0) \\ &\times G(y_0 \rightarrow y_1; \tau) G(y_1 \rightarrow y_2; \tau) \dots G(y_{L-1} \rightarrow y_L; \tau) e^{-S(Y)} \end{aligned} \quad (16)$$

The entire process of generating  $X$  starting with an initial  $x_0$ , followed by making  $Z$  from  $X$  and finally producing  $Y$  from  $Z$  is displayed below, in Figure 1. Now that we have both  $X$  and  $Y$ , we can attempt to sample from the target distributions.

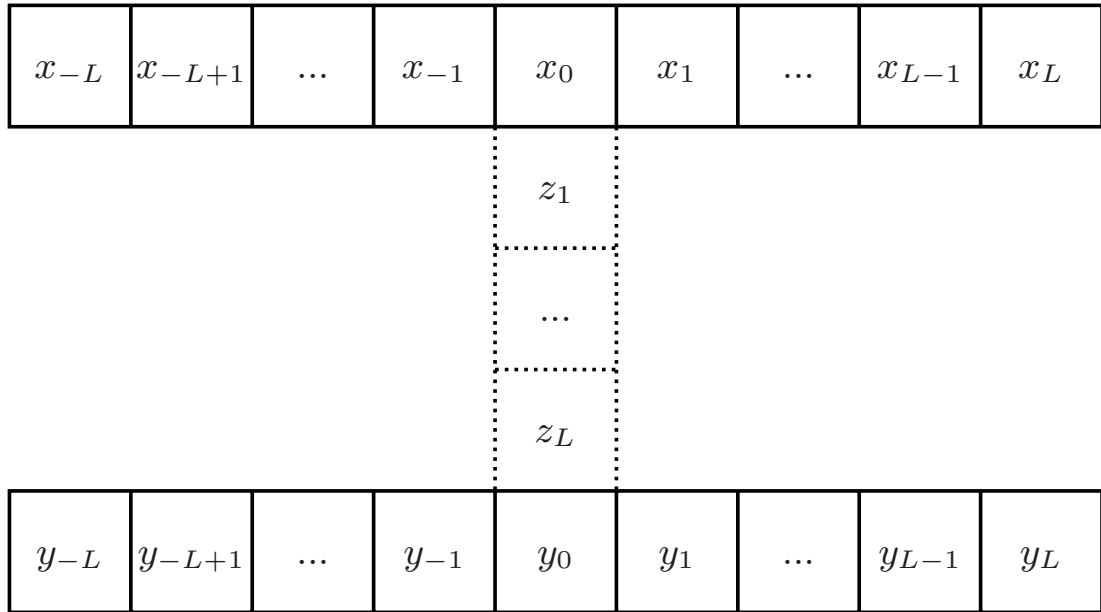


Figure 1: Visual representation of the process used to generate  $X$  and  $Y$  in our methodology. Path  $X$  is generated from  $x_0$ . Intermediate path  $Z$  is generated starting from  $x_0$  as well. Path  $Y$  is generated starting from  $z_L$ . All the moves are done using drift and diffusion. The paths are shown running in either vertical or horizontal direction for the sake of illustration. They can actually run in any direction and may even cross each other.

To perform our sampling, we must choose whether to accept  $Y$  or reject it and keep  $X$ . This is done by using the so-called *Metropolis decision*, written in the most general form as:

$$A(X \rightarrow Y) = \min \left\{ 1, \frac{\hat{\Pi}(Y)W(Y \rightarrow X)}{\hat{\Pi}(X)W(X \rightarrow Y)} \right\} \quad (17)$$

We know precisely what our target distributions,  $\hat{\Pi}(X)$  and  $\hat{\Pi}(Y)$ , are. In order to obtain a Metropolis decision however, we need to write down the proposal densities,  $W(X \rightarrow Y)$  and  $W(Y \rightarrow X)$ , associated with moving from  $X$  to  $Y$  and from  $Y$  to  $X$ , respectively. The proposal density for moving from  $X$  to  $Y$  can be written as:

$$\begin{aligned} W(X \rightarrow Y) = & G(x_0 \rightarrow z_1; \tau) G(z_1 \rightarrow z_2; \tau) \dots G(z_{L-1} \rightarrow z_L; \tau) G(z_L \rightarrow y_0; \tau) \\ & \times G(y_0 \rightarrow y_{-1}; \tau) G(y_{-1} \rightarrow y_{-2}; \tau) \dots G(y_{-L+1} \rightarrow y_{-L}; \tau) \\ & \times G(y_0 \rightarrow y_1; \tau) G(y_1 \rightarrow y_2; \tau) \dots G(y_{L-1} \rightarrow y_L; \tau) \end{aligned} \quad (18)$$

Similarly, the proposal density for moving from  $Y$  to  $X$  can be expressed as:

$$\begin{aligned} W(Y \rightarrow X) = & G(y_0 \rightarrow z_L; \tau) G(z_L \rightarrow z_{L-1}; \tau) \dots G(z_2 \rightarrow z_1; \tau) G(z_1 \rightarrow x_0; \tau) \\ & \times G(x_0 \rightarrow x_{-1}; \tau) G(x_{-1} \rightarrow x_{-2}; \tau) \dots G(x_{-L+1} \rightarrow x_{-L}; \tau) \\ & \times G(x_0 \rightarrow x_1; \tau) G(x_1 \rightarrow x_2; \tau) \dots G(x_{L-1} \rightarrow x_L; \tau) \end{aligned} \quad (19)$$

Substituting the expressions for the target distributions of  $X$  and  $Y$  along with their corresponding proposal densities into Equation (17), we can express the Metropolis decision in the following form:

$$A(X \rightarrow Y) = \min \left\{ 1, \frac{\Psi^2(y_0) G(y_0 \rightarrow z_L; \tau) G(z_L \rightarrow z_{L-1}; \tau) \dots G(z_2 \rightarrow z_1; \tau) G(z_1 \rightarrow x_0; \tau) e^{-S(Y)}}{\Psi^2(x_0) G(x_0 \rightarrow z_1; \tau) G(z_1 \rightarrow z_2; \tau) \dots G(z_{L-1} \rightarrow z_L; \tau) G(z_L \rightarrow y_0; \tau) e^{-S(X)}} \right\} \quad (20)$$

Although some of the Green's functions have cancelled upon substitution, thereby reducing the complexity of the Metropolis decision, some have remained. We can further reduce Equation (20) by assuming microscopic reversibility, which is valid only for small  $\tau^c$ ,

$$\Psi^2(x) G(x \rightarrow x'; \tau) = \Psi^2(x') G(x' \rightarrow x; \tau) \quad (21)$$

---

<sup>c</sup>The assumption of microscopic reversibility holds true only in the limit of zero time-step. This is analogous to our calculations of the ground-state energy and other properties: they also hold true only in the limit of vanishing time-step. We meet this criterion by extrapolating our results in the limit of zero time-step and producing results free of time-step bias.

Assuming microscopic reversibility along path  $Z$ , such that:

$$\begin{aligned}\Psi^2(x_0)G(x_0 \rightarrow z_1; \tau) &= \Psi^2(z_1)G(z_1 \rightarrow x_0; \tau) \\ \Psi^2(z_1)G(z_1 \rightarrow z_2; \tau) &= \Psi^2(z_2)G(z_2 \rightarrow z_1; \tau) \\ &\dots \\ \Psi^2(z_L)G(z_L \rightarrow y_0; \tau) &= \Psi^2(y_0)G(y_0 \rightarrow z_L; \tau)\end{aligned}$$

we derive the following expression:

$$\begin{aligned}\Psi^2(x_0)G(x_0 \rightarrow z_1; \tau)G(z_1 \rightarrow z_2; \tau)\dots G(z_{L-1} \rightarrow z_L; \tau)G(z_L \rightarrow y_0; \tau) = \\ \Psi^2(y_0)G(y_0 \rightarrow z_L; \tau)G(z_L \rightarrow z_{L-1}; \tau)\dots G(z_2 \rightarrow z_1; \tau)G(z_1 \rightarrow x_0; \tau)\end{aligned}\quad (22)$$

Substituting this result into Equation (20), we obtain the simplified expression for our Metropolis decision in its final form:

$$A(X \rightarrow Y) = \min \{1, e^{S(X)-S(Y)}\} \quad (23)$$

We employ this Metropolis decision to choose whether we want to accept  $Y$  or reject it and keep  $X$  instead. If  $Y$  is accepted, it is immediately re-labelled as  $X$  in order to be used in the next iteration of the algorithm. After the Metropolis decision, the algorithm generates a pure distribution,  $\Phi_0^2$ , at the middle of the path, and mixed distributions,  $\Psi\Phi_0$ , at the ends. This is shown in Figure 2.

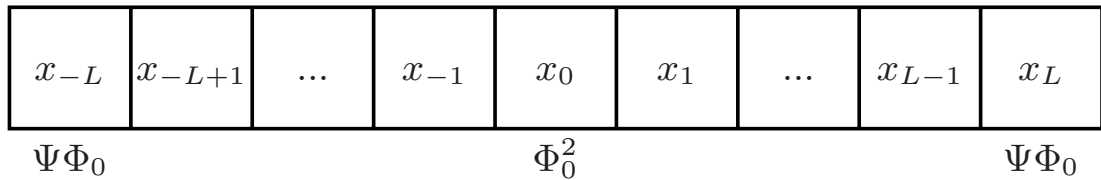


Figure 2: Schematic representation of an electron distribution along a path selected using the Metropolis decision. The algorithm generates a pure distribution,  $\Phi_0^2$ , at the middle of the path and mixed distributions,  $\Psi\Phi_0$ , at the ends.

When a new path is generated using drift and diffusion moves, the distribution

along this path is entirely variational,  $\Psi^2$ . The Metropolis decision is used to either accept the new path or reject it and keep the old path. It provides a statistical weight proportional to  $\Phi_0/\Psi$  to the ends of the chosen path, allowing one to sample from the mixed distribution,  $\Psi\Phi_0$ , at those locations. This mixed distribution is the solution to our modified Schrödinger equation and we can sample it to recover the fixed-node energy. Similarly, the Metropolis decision provides a statistical weight proportional to  $(\Phi_0/\Psi)^2$  to the middle of the chosen path, permitting one to sample from the pure distribution,  $\Phi_0^2$ , at that location. We sample the pure distribution to calculate various electronic properties that are represented by operators that do not commute with the Hamiltonian and described in detail in the following section. This process of providing statistical weights to the variational distribution was originally presented by Liu and co-workers.<sup>46</sup>

At this point, let us discuss the significance of the intermediate path  $Z$  that links paths  $X$  and  $Y$ . While its electron configurations do not appear in the final Metropolis decision, it nonetheless serves two vital purposes. First, it allows us to reduce the Metropolis decision to its final form, given by Equation (23). Second, it introduces sufficient separation in imaginary time between two paths,  $X$  and  $Y$ , to reduce the serial correlation of sampled properties whenever path  $Y$  is accepted. Finally, by setting the length of  $Z$  equal to that of  $X$  and  $Y$ , our algorithm has a single parameter that is readily optimized.

We demonstrated that by doing a sequence of small moves in imaginary time and evoking the Metropolis decision we can sample the mixed and pure distributions of electrons to determine the ground-state energy and properties, respectively. Our proposed solution has two inherent biases: it is only valid in the limit of zero time-step and infinite imaginary time. To remove these biases, we perform simulations for a range of different time-steps, each having a path that contains the electron configurations of different length. After sampling our properties, we extrapolate them

in the limit of zero-time step, thereby removing the biases by approaching the limit of zero time-step and infinite imaginary (simulation) time.

## 2.4 Fixed-node Approximation

In principle, our proposed method of solving the Schrödinger equation should yield the exact ground-state energy and electronic properties. However, upon its implementation, we run into the famous *fermion sign problem*.<sup>32</sup> The wave function of a multi-electron system must be antisymmetric (i.e., it must change sign) under the exchange of any two electrons with the same spin, albeit with an exception of some simple electronic systems such as the helium singlet. Furthermore, such a wave function can have exchange nodes: instances when the wave function is equal to zero for specific locations of electrons with the same spin. These conditions give rise to a nodal hypersurface: positive and negative regions within the wave function that are separated by nodes. The exact nodal hypersurface of a wave function can only be obtained from a complete solution of the Schrödinger equation, but some of its mathematical properties are known.<sup>6,14,30–32,47–50</sup>

The fact that we do not know the exact nodal hypersurface of a wave function *a priori* leads to a problem. In our PSQMC method, the mixed distribution,  $\Psi\Phi_0$ , must be positive definite everywhere. Whenever two electrons with the same spin exchange their positions,  $\Psi$  will change sign to preserve its antisymmetry, thereby making the mixed distribution negative. To ameliorate this problem, we introduce the *fixed-node approximation* by imposing the nodes of  $\Psi$  onto  $\Phi_0$ .<sup>6,14,30–32</sup> Therefore, whenever  $\Psi$  changes sign  $\Phi_0$  must also change sign, ensuring that the mixed distribution is positive definite everywhere. We denote the ground-state wave function within the fixed-node approximation by  $\Phi_{FN}$ . Since  $\Phi_{FN}$  has the same nodal structure as  $\Psi$ , which may not be exact, the accuracy of our method depends on the quality of the nodal structure of  $\Psi$ .



Using the variational principle, we can show that the fixed-node energy is an upper bound to the exact ground-state energy:<sup>39</sup>

$$E_{FN} = \frac{\langle \Phi_{FN} | \hat{H} | \Phi_{FN} \rangle}{\langle \Phi_{FN} | \Phi_{FN} \rangle} \geq E_0 \quad (24)$$

The energies reported by us in this work are the fixed-node ground-state energies, which are an upper bound to the exact ground-state energy.

## 2.5 Estimating the Energy

Our algorithm provides us with a straight-forward method to compute the fixed-node energy of a system in the ground-state. This is done by sampling the local energy from a mixed distribution of electrons and averaging it over the number of iterations of the algorithm. Formally, the energy calculated from a mixed distribution is given by the following equation (the derivation of which can be found in Appendix 2):

$$E_{FN} = \frac{1}{I} \sum_{i=1}^I E_{loc}(\mathbf{r}_i) \quad (25)$$

where  $I$  is the number of iterations of the algorithm and the local energy,  $E_{loc}$ , is calculated using Equation (4) from a mixed distribution of electrons  $\mathbf{r}_i$ .

Furthermore, our methodology has two independent locations<sup>d</sup> along the selected path where we can perform sampling from the mixed distribution. We can utilize both of them to provide a more accurate sample of the fixed-node energy by taking their average. Thus, our final expression for the fixed-node energy is:

$$E_{FN} = \frac{1}{I} \sum_{i=1}^I \frac{E_{loc}(x_{-L,i}) + E_{loc}(x_{L,i})}{2} \quad (26)$$

where  $I$  is the number of iterations of the algorithm and the local energy,  $E_{loc}$ , is

---

<sup>d</sup>They are independent provided there is a sufficient separation in imaginary time. This can be achieved by having a large-enough value of  $L$ . Conversely, having a small  $L$  would not guarantee their independence and the extreme case of  $L = 1$  would suggest that we are sampling from a distribution that is close to a variational distribution.

sampled from the selected path at two locations,  $x_{-L,i}$  and  $x_{L,i}$ , simultaneously. This technique provides us with a very good estimate of the fixed-node energy, as evident by our results.

## 2.6 Truncations

We do not employ a Jastrow function<sup>51</sup> to account for the electron-electron and electron-nucleus cusp conditions ( $1/r$  type singularities formed when two Coulomb particles come close to each other). To avoid unreasonable drift and diffusion moves and absurd values of local energies, we introduce truncations that depend on the time-step, so that their associated biases vanish as the time-step approaches zero.<sup>52</sup> In addition, time-step bias present in the Green's function (Equation (8)) and bias from the assumption of microscopic reversibility also vanish in that limit. The respective truncations are given by:

$$F_\alpha = \begin{cases} F_\alpha & \text{if } |F_\alpha| \leq 1/\tau \\ \text{sign}[1/\tau, F_\alpha] & \text{otherwise} \end{cases} \quad (27)$$

and

$$(E_{loc} - E_T) = \begin{cases} (E_{loc} - E_T) & \text{if } |E_{loc} - E_T| \leq 1/\tau \\ \text{sign}[1/\tau, (E_{loc} - E_T)] & \text{otherwise} \end{cases} \quad (28)$$

where the Greek letter subscript  $\alpha$  represents the Cartesian coordinates  $x$ ,  $y$  or  $z$ .  $F_\alpha$  is a component of the drift velocity given by Equation (3) and  $\tau$  is the time-step.  $E_T$  is the trial energy cutoff that was set to  $-8.0 E_h$  for lithium hydride,  $-76.0 E_h$  for water and  $-78.0 E_h$  for ethylene. The energy cutoff is not sensitive to the choice of  $E_T$ . In practice these cutoffs are rarely evoked.

### 3 Electronic Properties

Calculating properties of experimental importance, such as the dipole moment, is extremely straight-forward in our methodology. This is due to our ability to generate and sample the pure distribution of electrons at the middle of the selected path. We can sample any property, which commutes with the position operator, from the pure distribution using the following formula:

$$\langle \hat{A} \rangle = \frac{1}{I} \sum_{i=1}^I \hat{A}(x_{0,i}) \quad (29)$$

where  $I$  is the number of iterations of the algorithm and  $\hat{A}$  is the property of interest sampled from the pure distribution located at  $x_{0,i}$ . The properties explored by us in this work are described in detail below. There is a multitude of other properties that can be calculated from the pure distribution using our algorithm, which were not explored due to time constraints.

#### 3.1 Electric Moments

Consider a system of charged particles that is overall electrically neutral, but contains within itself a group of positively-charged particles at one end and a group of negatively-charged particles at the opposite end. Such a system is an *electric dipole* that has an *electric dipole moment*, a quantity that measures the degree of separation of positive and negative charges inside the system. The dipole moment is said to be *permanent* if it arises from a static configuration of charged particles in the system, or *induced* if it stems from an applied external electric field.

An asymmetrical molecule that has a permanent dipole moment due to the arrangement of its atoms is called a *polar* molecule. Polar molecules can interact with each-other via attraction between their oppositely-charged regions. This effect is called the *dipole-dipole interaction*. Water molecule is a good example of this.

The oxygen atom in the water molecule is more negative than the hydrogen atoms. Thus, individual water molecules can form a so-called hydrogen bond with each other through the dipole-dipole interaction between the oxygen atom on one molecule and a hydrogen atom on another molecule. This effect gives rise to many unique properties of water.

Mathematically, the dipole moment arises from an expansion of the electric potential for a system of charged particles. This expansion produces other important terms besides the dipole moment. Using Buckingham's notation<sup>53</sup>, we can write down equations that define the dipole ( $\boldsymbol{\mu}$ ), quadrupole ( $\boldsymbol{\Theta}$ ) and octupole ( $\boldsymbol{\Omega}$ ) moments in the following form:

$$\mu_{\alpha} = \sum_i q_i r_{i\alpha} \quad (30)$$

$$\Theta_{\alpha\beta} = \frac{1}{2} \sum_i q_i (3r_{i\alpha}r_{i\beta} - r_i^2 \delta_{\alpha\beta}) \quad (31)$$

$$\Omega_{\alpha\beta\gamma} = \frac{1}{2} \sum_i q_i (5r_{i\alpha}r_{i\beta}r_{i\gamma} - r_i^2 (r_{i\alpha}\delta_{\beta\gamma} + r_{i\beta}\delta_{\gamma\alpha} + r_{i\gamma}\delta_{\alpha\beta})) \quad (32)$$

Here, the Greek letter subscripts  $\alpha$ ,  $\beta$  and  $\gamma$  represent the Cartesian coordinates  $x$ ,  $y$  or  $z$ . The summations occur over all charged particles in the molecule, both electrons and nuclei:  $q_i$  represents the charge of the  $i^{th}$  particle and  $r_i$  denotes its position with respect to the centre of mass of the molecule. Furthermore,  $\Theta_{\alpha\beta}$  and  $\Omega_{\alpha\beta\gamma}$  are symmetric under the interchange of their indices.

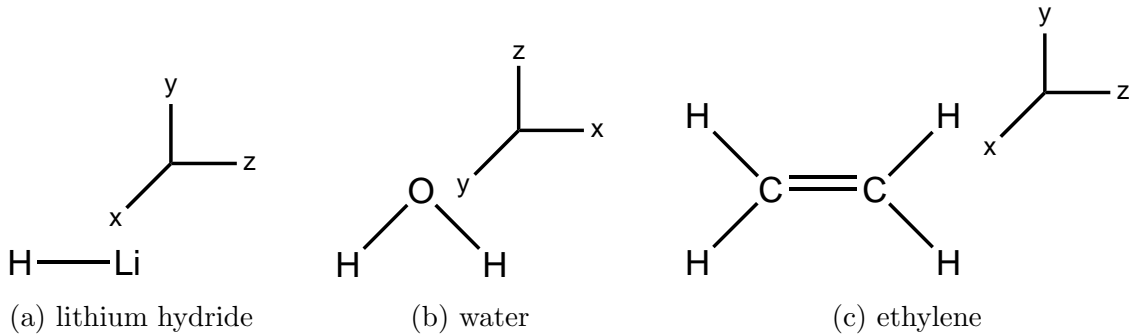


Figure 3: Illustration of geometries for (a) lithium hydride, (b) water and (c) ethylene.

Due to the symmetry of lithium hydride, the only non-zero component of the dipole moment is  $\mu_z$ . The following tensor properties<sup>54</sup> for  $\Theta$  and  $\Omega$  apply to lithium hydride:

$$\Theta_{xx} = \Theta_{yy}, \Theta_{zz} = -(\Theta_{xx} + \Theta_{yy}) = -2\Theta_{xx} \quad (33)$$

$$\Omega_{xxz} = \Omega_{yyz}, \Omega_{zzz} = -(\Omega_{xxz} + \Omega_{yyz}) = -2\Omega_{xxz} \quad (34)$$

There is a similar set of rules that apply to the water molecule. Due to its symmetry,  $\mu_z$  is the only non-zero component of the dipole moment. The following tensor properties<sup>54</sup> for  $\Theta$  and  $\Omega$  apply to the water molecule:

$$\Theta_{zz} = -(\Theta_{xx} + \Theta_{yy}) \quad (35)$$

$$\Omega_{zzz} = -(\Omega_{xxz} + \Omega_{yyz}) \quad (36)$$

Due to the symmetry of the ethylene molecule, its dipole and octupole moments are both zero. The following tensor property<sup>54</sup> for  $\Theta$  applies to the ethylene molecule:

$$\Theta_{zz} = -(\Theta_{xx} + \Theta_{yy}) \quad (37)$$

In our work, the components of all electric moments are calculated independently. Therefore, the aforementioned symmetry requirements can be useful in various important ways. We can confirm the validity of our results by checking to see if they obey the symmetry rules. Furthermore, we can improve the quality of our results by taking a variance-weighted average of symmetry-equivalent components, where applicable. For example, our result for  $\Theta_{xx}$  of lithium hydride is obtained from a variance-weighted average of  $\Theta_{xx}$  and  $\Theta_{yy}$ .

### 3.2 Diamagnetic Shielding and Susceptibility

Another set of properties that we examine are the diamagnetic shielding and diamagnetic susceptibility, expressed by Equations (38) and (39), respectively. We do not

calculate their paramagnetic contributions since we do not extract the excited-state energies of molecules in this work.

$$1/r_A = \sum_i^n \frac{1}{|r_i - R_A|} \quad (38)$$

$$r^2 = \sum_i^n |r_i - R_{CM}|^2 \quad (39)$$

Here,  $n$  is the number of electrons in the molecule.  $R_A$  is the location of the nucleus  $A$  for which the diamagnetic shielding is being calculated.  $R_{CM}$  represents the location of the centre of mass.

### 3.3 Electric Fields and Electric Field Gradients

In a molecule at equilibrium geometry, the total electric field felt by each nucleus is zero. We can examine the quality of our results and the accuracy of chosen geometry by calculating the electric fields and observing any significant deviations from zero. In addition, we study electric field gradients that measure the rate at which the electric field changes at a location of a specific nucleus. We can compare them to known quantities and report previously-unknown ones. One can use electric field gradients to calculate the quadrupole coupling constant, a useful experimental quantity, by combining it with a nuclear quadrupole moment.<sup>55</sup>

The expressions for the intramolecular electric fields and electric field gradients are given by Equations (40) and (41), respectively.

$$q\alpha/r_A^3 = - \sum_{\substack{i=1 \\ i \neq A}}^N \frac{Z_i (R_{i\alpha} - R_{A\alpha})}{|R_i - R_A|^3} + \sum_{\substack{i=1 \\ \sim}}^n \frac{(r_{i\alpha} - R_{A\alpha})}{|r_i - R_A|^3} \quad (40)$$

$$\begin{aligned}
q_A(\alpha\beta) = & - \sum_{\substack{i=1 \\ i \neq A}}^N \frac{Z_i [3(R_{i\alpha} - R_{A\alpha})(R_{i\beta} - R_{A\beta}) - \delta_{\alpha\beta}|R_i - R_A|^2]}{|R_i - R_A|^5} \\
& + \sum_{\substack{i=1 \\ \sim}}^n \frac{[3(r_{i\alpha} - R_{A\alpha})(r_{i\beta} - R_{A\beta}) - \delta_{\alpha\beta}|r_i - R_A|^2]}{|r_i - R_A|^5}
\end{aligned} \tag{41}$$

Again, the Greek letter subscripts  $\alpha$  and  $\beta$  represent the Cartesian coordinates  $x$ ,  $y$  or  $z$ . Variable  $A$  represents the nucleus for which the quantities are being calculated.  $N$  and  $n$  are the number of nuclei and electrons in the molecule, respectively. The  $\sim$  symbol on the bottom of the sum indicates that the  $1/r$  type singularity has been replaced with an appropriate truncation, described as follows.

When we calculate the electric fields, we replace  $1/r^3$  by  $5/2\epsilon^3$  whenever  $r \leq \epsilon$ , where  $r = |r_i - R_A|$  and  $\epsilon = 2\tau^{1/3}$ , a small  $\tau$ -dependent number. Similarly, when we calculate the electric field gradients, we replace  $1/r^5$  by  $7/2\epsilon^5$  whenever  $r \leq \epsilon$ . These truncations reduce the variance of simulated properties by several orders of magnitude.<sup>56</sup> Furthermore, they are designed to vanish in the  $\tau \rightarrow 0$  limit, so that they do not introduce any bias into the extrapolated values.

### 3.4 Polarizabilities

Electric polarizability describes the affinity of a charge distribution to be distorted by an applied electric field. When a charge distribution experiences an external electric field, an induced dipole moment is created, whose magnitude is proportional to the polarizability and the size of the field. Polarizability sheds light on the understanding of intermolecular interactions between nonpolar atoms or molecules and polar ones that are either charged or have a permanent dipole moment. The appearance of the induced dipole moment in nonpolar molecules allows them to interact with their neighbours via London dispersion forces (intermolecular forces that arise due to instantaneous interactions between induced dipole moments). We calculate the components of the static  $\alpha$  polarizability using the same method we employed previously

in the context of reptation quantum Monte Carlo.<sup>57</sup>

The components of the static  $\alpha$  polarizability are given by the following equation:

$$\alpha_{\alpha\beta} = \tau \text{cum}(\{\mu_\alpha\}, \mu_\beta) \quad (42)$$

where the Greek letter subscripts  $\alpha$  and  $\beta$  represent the Cartesian coordinates  $x$ ,  $y$  or  $z$  and  $\tau$  is the time-step. The *cumulant* is defined as:

$$\text{cum}(A, B) = \overline{([A - \bar{A}] \cdot [B - \bar{B}])} \quad (43)$$

where the bar represents an average value. The dipole moment enclosed in curly brackets is calculated using:

$$\{\mu_i\} = \sum_{l=-L}^L \mu_i(x_l) + \mu_i(x_0) \quad (44)$$

where the individual dipole moments  $\mu_i$  are sampled from electron configurations  $x_l$  along the accepted path  $X$ , which has a total length of  $2L + 1$  (refer to Figure 2 for details).

It is important to observe that the above formalism is free from the finite-field approximation, which is sometimes used to estimate the energy change caused by an external field. This point is critically important in the context of stochastic estimation of polarizabilities, where such an approach may be confounded by random fluctuations of the simulated energy.



## 4 Results and Discussion

We applied our newly-developed pure-sampling quantum Monte Carlo (PSQMC) method to lithium hydride, water and ethylene molecules, and calculated their ground-state energies and electronic properties within the fixed-node approximation.<sup>35</sup> A single-determinant importance sampling function with a QZ4P STO basis set was generated using Amsterdam Density Functional (ADF)<sup>58–60</sup> software for each molecule. The calculations were carried out using experimental geometries at various time-steps and path-lengths (these details are specified below for each molecule in its respective section). As we decreased the time-step, we increased the path-length using Equation (45) for lithium hydride and water, and Equation (46) for ethylene.

$$L(\tau) = L_0 \left( \frac{\tau_0}{\tau} \right)^{3/2} \quad (45)$$

$$L(\tau) = L_0 \left( \frac{\tau_0}{\tau} \right) \quad (46)$$

where  $L_0$  is the initial (smallest) path-length and  $\tau_0$  is the initial (largest) time-step. Note that the path-length ( $L$ ) increases as the time-step ( $\tau$ ) decreases from the  $\tau_0$  value, approaching infinity as  $\tau$  approaches zero. This gives rise to nearly constant error bars on the properties for the entire range of time-steps.

Figures 4, 7 and 9 display plots of the ground-state energy of lithium hydride, water and ethylene molecules, respectively, as a function of time-step and path-length parameter ( $L_0$ ). These plots were generated as follows: As  $L_0$  is systematically increased, a regression model is chosen such that its  $\tau = 0$  intercept (determined using variance-weighted regression) is as consistent as possible with that of the previous  $L_0$ . As the path-length increases and passes a critical value,  $L_0 = 41$  for lithium hydride,  $L_0 = 151$  for water and  $L_0 = 60$  for ethylene, the ground-state energy converges to a consistent value. Any further increase of the path-length beyond a critical value yields equivalent results, albeit with an increased computational cost associated with

an increased number of electron configurations being generated along the paths.

While performing multiple runs with different path-length parameters that extend beyond a critical value may seem redundant, the observed convergence instills confidence in our choice of  $L_0$ . Furthermore, variance-weighted averaging of converged results provides us with an accurate estimate of the property that is free from time-step and regression-model biases. In addition, the convergence exhibited by the ground-state energy is also demonstrated by all ground-state electronic properties that we report in this work, which are calculated using the same approach. When calculating properties, the path-lengths at which the ground-state energy did not converge are discarded. To our knowledge, this feature of our approach is not present in any other quantum Monte Carlo method.

## 4.1 Application to Lithium Hydride

Table 1: Ground-state energy and dipole moment of lithium hydride in the limit of zero time-step ( $\tau$ ) and as a function of path-length parameter ( $L_0$ ). All entries are in atomic units (a.u.).

$L_0$	$E_0$	$\mu_z$
11	$-8.06090(64)^\dagger$	$2.3498(21)^\ddagger$
21	$-8.06643(83)^\dagger$	$2.3515(36)^\ddagger$
31	$-8.06896(31)^\dagger$	$2.3501(16)^\ddagger$
41	$-8.07011(96)$	$2.3373(63)^\dagger$
51	$-8.07060(57)$	$2.3433(11)^\dagger$
101	$-8.07043(119)$	$2.3299(30)^\dagger$
151	$-8.07093(90)$	$2.3279(17)^\dagger$
201	$-8.07085(38)$	$2.3096(42)$
251	$-8.07033(140)$	$2.3058(39)$
301	$-8.07074(152)$	$2.3055(33)$
Average*	$-8.07070(27)$	$2.3067(22)$

<sup>†</sup> Not included in the weighted average due to a lack of convergence with the rest of the data.

<sup>‡</sup> Not included in the weighted average because the ground-state energy at this path-length did not converge.

\* Variance-weighted average of the property for all included path-lengths. These properties and their corresponding literature values are also quoted in Table 2.

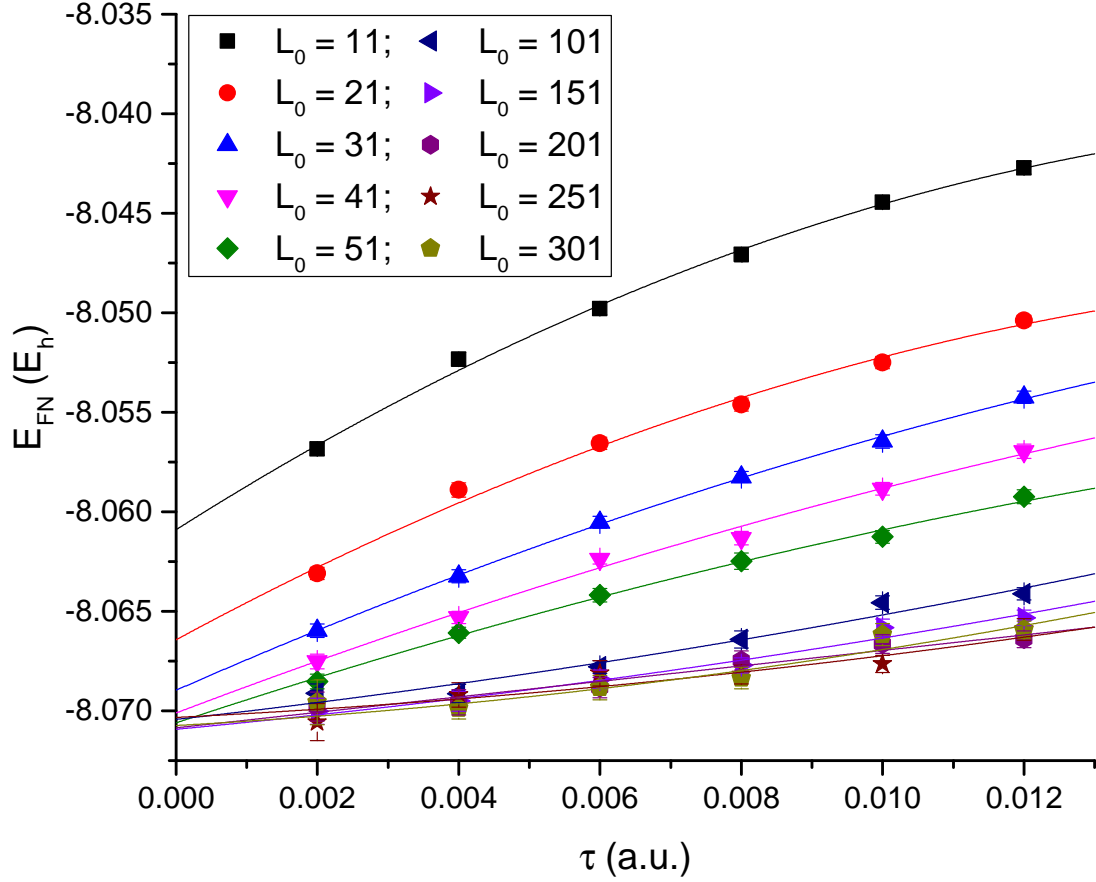


Figure 4: Ground-state energy of lithium hydride as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (45), where  $\tau_0 = 0.012$  a.u. Convergence occurs for  $L_0 = 41 \dots 301$ [50].

For each path-length parameter in our calculations for lithium hydride, we performed 100 independent runs with 20 000 iterations per run and with an additional 5 000 iterations thrown out for the purpose of equilibration. We used experimental geometry from Reference 61 in our calculations. Table 2 contains our results for lithium hydride. The ground-state energy was calculated by variance-weighted averaging of converged results from eight separate values of the path-length parameter,  $L_0 = 41 \dots 301$ [50]; see Figure 4 and Table 1 for details. The ground-state energy is in excellent agreement with the literature and agrees exactly with Nemec *et al.*,<sup>62</sup> who used an identical trial wave function.

Similarly, the dipole moment was obtained by variance-weighted averaging of

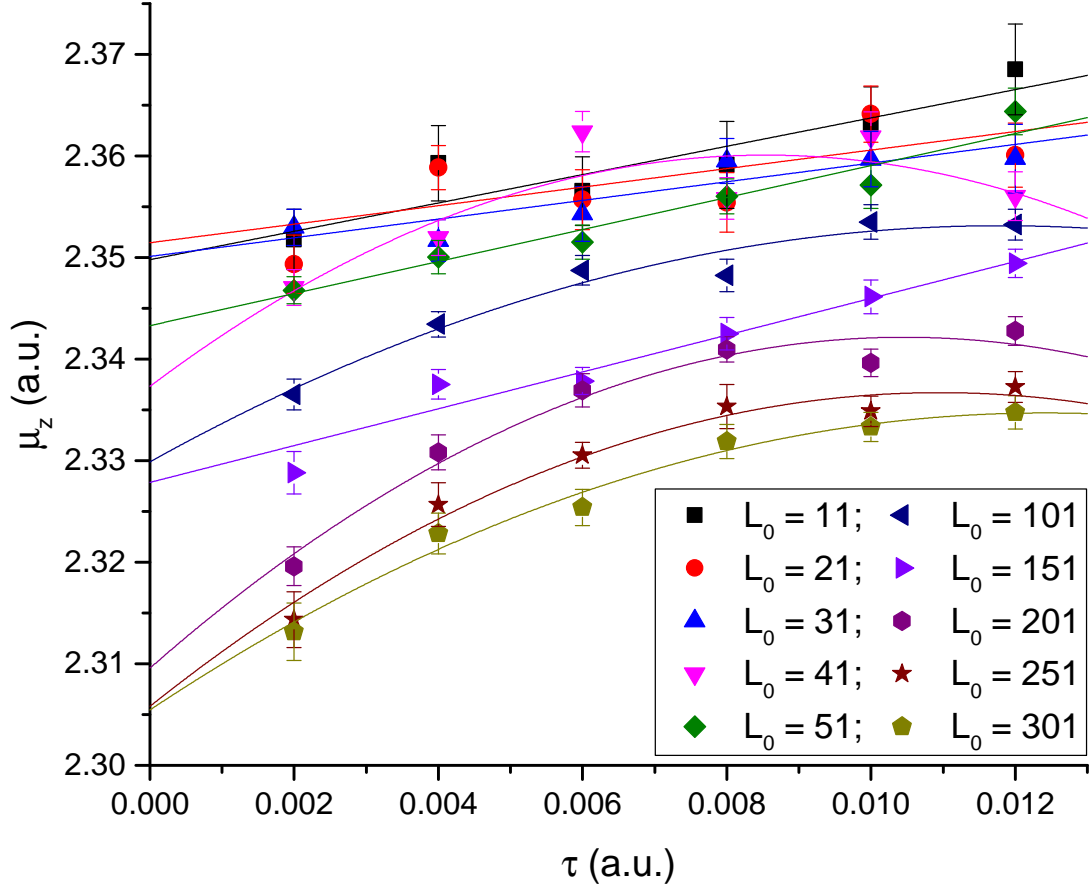


Figure 5: Dipole moment of lithium hydride as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (45), where  $\tau_0 = 0.012$  a.u. Convergence occurs for  $L_0 = 201 \dots 301$ [50].

converged results from three distinct values of the path-length parameter,  $L_0 = 201 \dots 301$ [50]; see Figure 5 and Table 1 for details. The dipole moment agrees with the theoretical and experimentally-determined literature values. The figures illustrating the convergence of the remaining electronic properties of lithium hydride are shown in Appendix 3. The quadrupole moment, octupole moment, diamagnetic shielding and diamagnetic susceptibility are in good agreement with their literature counterparts. The electric fields agree with their theoretical values, albeit with an exception of the  $z$  component on the lithium nucleus, which may reflect the fact that we used an experimental geometry rather than a geometry that minimizes the Hartree-Fock energy of the importance sampling function. The electric field gradients on the

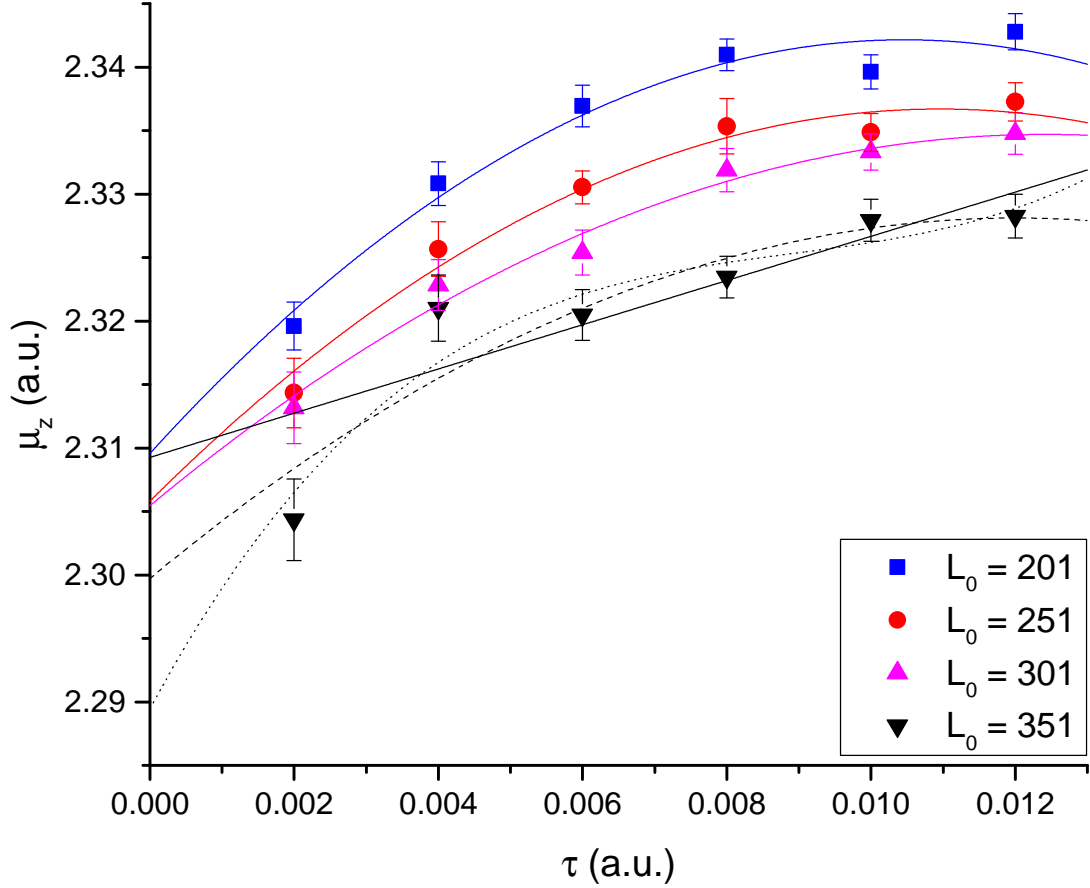


Figure 6: Dipole moment of lithium hydride as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (45), where  $\tau_0 = 0.012$  a.u. Convergence occurs for  $L_0 = 201 \dots 301$ [50]. No regression model can adequately fit the data for  $L_0 = 351$ : — linear fit; - - - quadratic fit;  $\cdots$  cubic fit.

nuclei are in excellent agreement with both theoretical and experimental literature values. Finally, our static  $\alpha$  polarizability is in excellent agreement with high-level theory, improving upon our previous Metropolis-Hastings reptation quantum Monte Carlo (RQMC-MH) results.<sup>57</sup>

There is a limit to how much one can increase  $L_0$  past its critical value. Beyond that limit, no regression model can be found to adequately fit the data. We attribute this instability to doing *Metropolis branching*<sup>7</sup> (using the Metropolis decision to accept a path and sample properties from its marginal distribution) with branching factors accumulated over a path that is too long. This is analogous to doing weight

branching in diffusion Monte Carlo with too many accumulated branching factors.<sup>63</sup> An example is  $L_0 = 351$  for calculating the dipole moment of lithium hydride. Here,  $L(\tau = 0.002) = 5159$  and the Metropolis decision is based upon accumulating 10 320 branching factors; see Equations (13) and (45). The resulting numerical instability is obvious upon inspection of Figure 6, where this data cannot be adequately fit by variance-weighted regression.

Table 2: Ground-state energy and electronic properties of lithium hydride. All entries are in atomic units (a.u.).

Property	PSQMC	Bounce <sup>a</sup>	RQMC-MH <sup>b</sup>	Literature
$E_0$	-8.0707(3)	-8.073(2)	-8.0701(4)	-8.0704(2) <sup>c</sup> , -8.070553(5) <sup>d</sup>
$\mu_z$	2.307(2)	2.35(3)	2.291(6)	2.2939 <sup>e</sup> , 2.3140 <sup>f</sup> , 2.314(1) <sup>g</sup>
$\Theta_{xx}$	1.558(5)	1.67(5)	1.55(2)	1.5485 <sup>e</sup> , 1.554 <sup>h</sup>
$\Theta_{zz}$	-3.11(1)	-3.33(8)	-3.10(3)	-3.097 <sup>e</sup> , -3.108 <sup>h</sup>
$\Omega_{xxz}$	-2.80(2)	-3.72(29)	-2.75(7)	-2.84(2) <sup>i</sup> , -2.9031 <sup>j</sup>
$\Omega_{zzz}$	5.61(2)	7.35(31)	5.50(14)	5.68(2) <sup>i</sup> , 5.8062 <sup>j</sup>
$\langle 1/r_{Li} \rangle_e$	6.085(3)	6.078(2)	5.912(4)	6.08 <sup>k</sup>
$\langle 1/r_H \rangle_e$	2.2366(8)	2.24(1)	2.231(2)	2.73 <sup>k</sup>
$\langle r^2 \rangle_e$	22.300(6)	22.62(16)	22.28(4)	22.596 <sup>l</sup>
$\langle qx/r_{Li}^3 \rangle$	-0.002(2)	0.002(4)	-0.0002(3)	0.0
$\langle qy/r_{Li}^3 \rangle$	-0.002(2)	-0.002(4)	-0.0001(3)	0.0
$\langle qz/r_{Li}^3 \rangle$	0.026(2)	0.011(6)	0.0010(8)	0.0
$\langle qx/r_H^3 \rangle$	-0.0010(8)	-0.008(5)	0.0002(4)	0.0
$\langle qy/r_H^3 \rangle$	-0.0006(4)	0.002(2)	-0.0003(4)	0.0
$\langle qz/r_H^3 \rangle$	0.003(1)	0.003(8)	0.018(3)	0.0
$q_{Li}(xx)$	0.021(3)	0.019(7)	0.0202(5)	0.02017 <sup>m</sup> , 0.0204(1) <sup>n</sup>
$q_{Li}(zz)$	-0.040(3)	-0.03(3)	-0.0404(7)	-0.04034 <sup>m</sup> , -0.0408(1) <sup>n</sup>
$q_H(xx)$	-0.025(1)	-0.025(4)	-0.024(2)	-0.02611 <sup>m</sup> , -0.0246(1) <sup>n</sup>
$q_H(zz)$	0.052(2)	0.052(6)	0.050(4)	0.05222 <sup>m</sup> , 0.0491(1) <sup>n</sup>
$\alpha_{xx}$	29.6(2)	—	27.92(4)	29.57 <sup>e</sup> , 30.9(4) <sup>i</sup> , 29.76 <sup>j</sup>
$\alpha_{zz}$	24.5(7)	—	23.6(2)	25.79 <sup>e</sup> , 24.6(4) <sup>i</sup> , 26.36 <sup>j</sup>

<sup>a</sup> Bounce quantum Monte Carlo, Reference 64.<sup>b</sup> Reptation quantum Monte Carlo, Reference 57.<sup>c</sup> Diffusion quantum Monte Carlo, Reference 62.<sup>d</sup> Explicitly correlated Gaussian (ECG), Reference 65.<sup>e</sup> Explicitly correlated coupled cluster (CCSD(T)-R12), Reference 66. The value for  $\Theta_{xx}$  is calculated from  $\Theta_{zz}$ .<sup>f</sup> Explicitly correlated Gaussian (ECG), Reference 67.<sup>g</sup> Experiment, Reference 68.<sup>h</sup> Doubly substituted coupled cluster (CCD/BO), Reference 69. The value for  $\Theta_{xx}$  is calculated from  $\Theta_{zz}$ .<sup>i</sup> Diffusion quantum Monte Carlo, Reference 56. The value for  $\Omega_{xxz}$  is calculated from  $\Omega_{zzz}$ .<sup>j</sup> MC SCF, Reference 70. The value for  $\Omega_{xxz}$  is calculated from  $\Omega_{zzz}$ .<sup>k</sup> SCF, Reference 71.<sup>l</sup> SCF, Reference 72.<sup>m</sup> SCF, Reference 73.<sup>n</sup> Experiment, References 68 and 74.

## 4.2 Application to Water

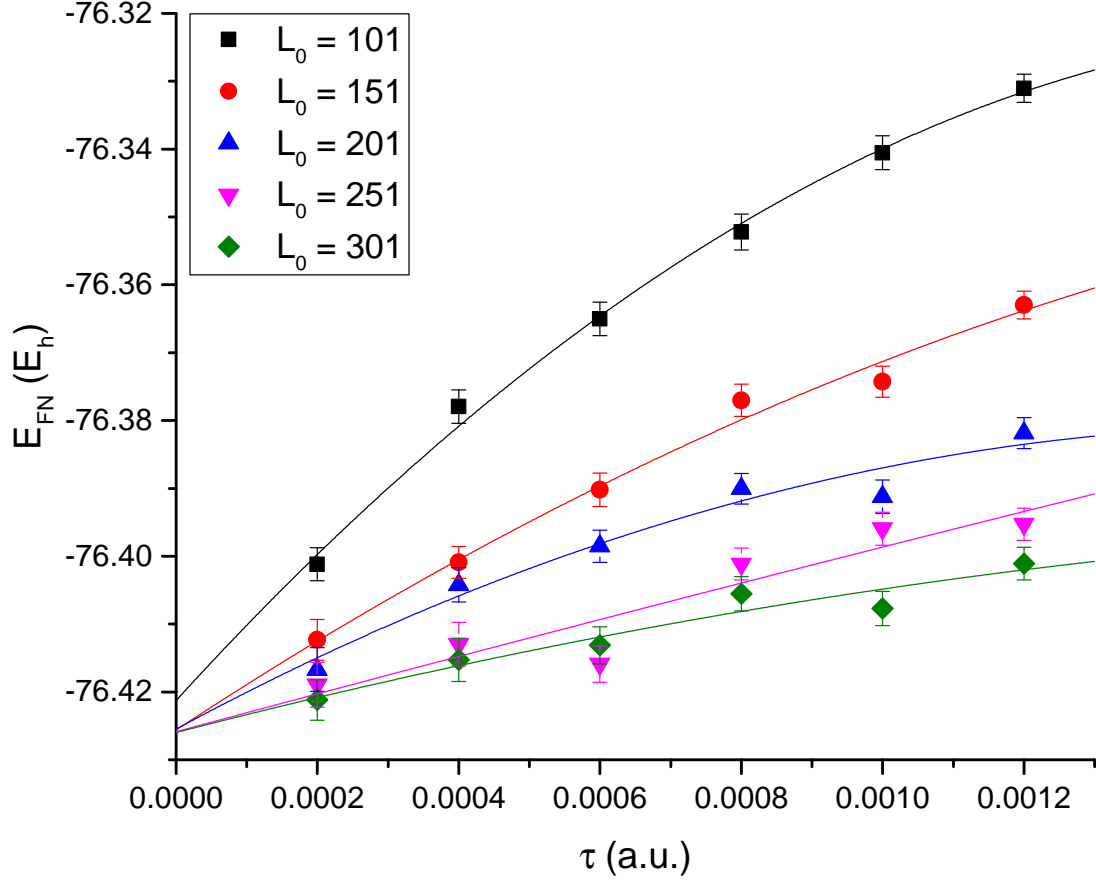


Figure 7: Ground-state energy of water as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (45), where  $\tau_0 = 0.0012$  a.u. Convergence occurs for  $L_0 = 151 \dots 301$ [50].

In our calculations for water, we performed 100 independent runs with 5 000 iterations per run and with an additional 1 000 iterations thrown out for the purpose of equilibration for each path-length parameter. We used experimental geometry from Reference 75. Our results for water are enclosed in Table 5. The ground-state energy and dipole moment were calculated by variance-weighted averaging of converged results from four separate values of the path-length parameter,  $L_0 = 151 \dots 301$ [50]; see Figures 7 and 8 and Table 3 for details. The figures illustrating the convergence of the remaining electronic properties of water are shown in Appendix 4. The ground-state energy is in excellent agreement with single-determinant literature and agrees exactly



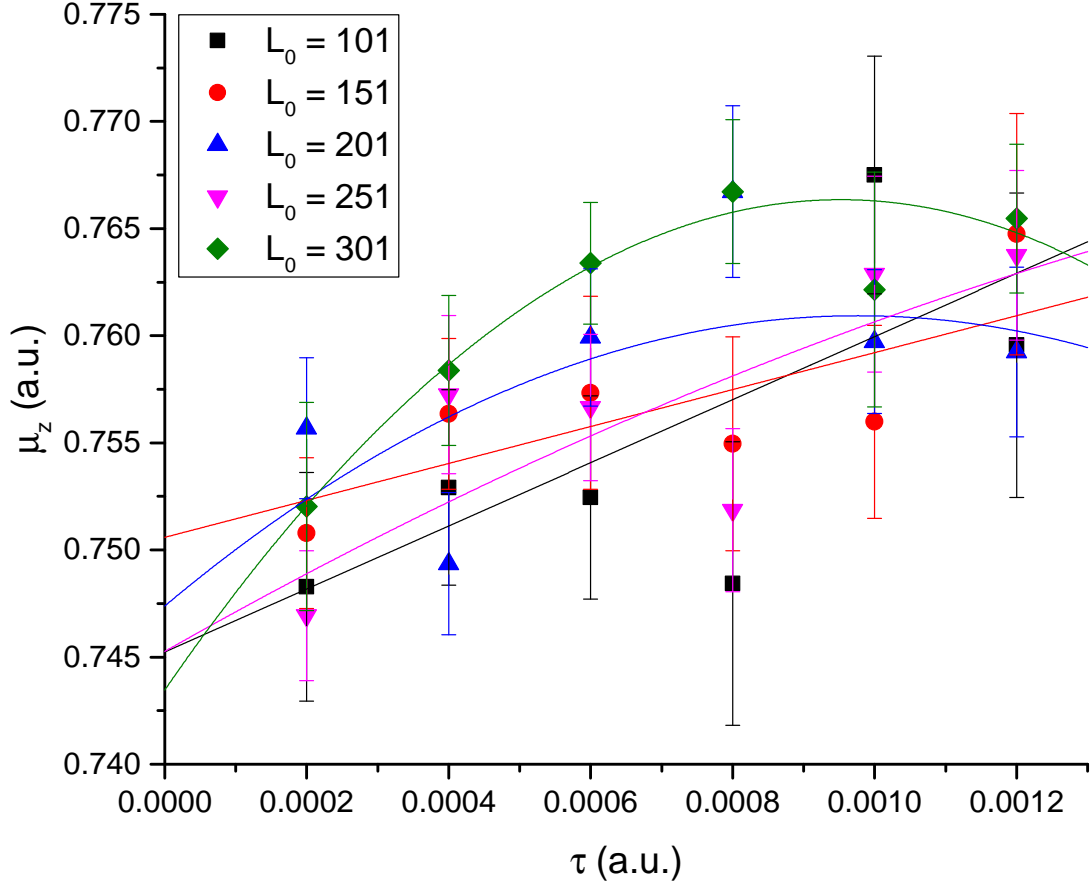


Figure 8: Dipole moment of water as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (45), where  $\tau_0 = 0.0012$  a.u. Convergence occurs for  $L_0 = 151 \dots 301$ [50].

with Nemec *et al.*,<sup>62</sup> who used an identical importance sampling function. When compared to methods that use a multi-reference approach, the energy is slightly more positive (our method guarantees an upper bound on the true ground-state energy). Nonetheless, this can be easily ameliorated by using multi-determinants, albeit with an increase in computational cost.<sup>76</sup>

Dipole, quadrupole and octupole moments, diamagnetic shielding and diamagnetic susceptibility are in agreement with the literature. There is a nominal deviation in the electric fields on the oxygen nucleus, which can be explained by our use of an experimental geometry rather than a geometry that minimizes the Hartree-Fock energy of the importance sampling function. The electric field gradients on the nu-

Table 3: Ground-state energy and dipole moment of water in the limit of zero time-step ( $\tau$ ) and as a function of path-length parameter ( $L_0$ ). All entries are in atomic units (a.u.).

$L_0$	$E_0$	$\mu_z$
101	$-76.4213(37)^\dagger$	$0.7452(49)^\ddagger$
151	$-76.4256(53)$	$0.7506(25)$
201	$-76.4255(66)$	$0.7474(100)$
251	$-76.4259(99)$	$0.7453(78)$
301	$-76.4260(50)$	$0.7435(38)$
Average*	$-76.4257(30)$	$0.7482(20)$

<sup>†</sup> Not included in the weighted average due to a lack of convergence with the rest of the data.

<sup>‡</sup> Not included in the weighted average because the ground-state energy at this path-length did not converge.

\* Variance-weighted average of the property for all included path-lengths. These properties and their corresponding literature values are also quoted in Table 5.

clei are in good agreement with both theoretical and experimental literature values.

Lastly, the static  $\alpha$  polarizability is in excellent agreement with high-level theory.

In several cases, our previous results obtained using reptation quantum Monte Carlo (RQMC) underestimate both PSQMC and corresponding literature values for lithium hydride and water. When it comes to pure-sampling of electronic properties for these molecules, PSQMC is generally more accurate and sometimes more precise than RQMC. However, it would be incorrect to conclude that pure-sampling done by PSQMC is inherently more accurate than by RQMC. This is still an open question. The path-length parameter has been optimized for PSQMC, while the chosen algorithmic parameters for RQMC may not have been ideal.

Table 4 shows additional literature results for static  $\alpha$  polarizability (third column) calculated with different coupled cluster methods. In our opinion, these results are unreliable and thus were not included in Table 5. Joshi *et al.*<sup>77</sup> report their ECCSD results (labeled by <sup>a</sup> in Table 4) for  $\alpha_{yy}$  and  $\alpha_{zz}$  as having the same value (within 0.01 a.u.). This cannot be the case due to the symmetry of the water molecule. Furthermore, according to our and literature results,  $\alpha_{zz}$  should be larger than  $\alpha_{yy}$  by

Table 4: Additional literature results for water polarizabilities. All entries are in atomic units (a.u.).

Property	This Work	Questionable Literature	Reliable Literature
$\alpha_{xx}$	9.7(3)	9.84 <sup>a</sup> , 9.353 <sup>b</sup>	9.944 <sup>c</sup> , 9.88 <sup>d</sup> , 9.74 <sup>e</sup> , 9.18 <sup>f</sup>
$\alpha_{yy}$	8.8(3)	9.14 <sup>a</sup> , 9.978 <sup>b</sup>	9.052 <sup>c</sup> , 8.92 <sup>d</sup> , 9.15 <sup>e</sup> , 7.90 <sup>f</sup>
$\alpha_{zz}$	9.4(1)	9.15 <sup>a</sup> , 9.613 <sup>b</sup>	9.398 <sup>c</sup> , 9.52 <sup>d</sup> , 9.36 <sup>e</sup> , 8.52 <sup>f</sup>

<sup>a</sup> Explicitly correlated coupled cluster (ECCSD), Reference 77.

<sup>b</sup> CCSD, Reference 78.

<sup>c</sup> CCSD, Reference 79.

<sup>d</sup> DFT, Reference 80.

<sup>e</sup> MC SCF, Reference 81.

<sup>f</sup> SCF, Reference 82.

a minimum of 0.3 a.u. It appears that  $\alpha_{zz}$  reported by Joshi *et al.* is under-estimated and  $\alpha_{yy}$  could be slightly over-estimated. Osted *et al.*<sup>78</sup> reports their CCSD results (labeled by <sup>b</sup> in Table 4) as  $\alpha_{yy} > \alpha_{zz} > \alpha_{xx}$ , when ordered by magnitude. Yet from our results and the literature, the magnitude of the polarizabilities should be ordered as  $\alpha_{xx} > \alpha_{zz} > \alpha_{yy}$ . According to Table IV in the work by Jensen *et al.*<sup>80</sup>, the dynamic polarizabilities of water molecule are higher than the static polarizabilities and increase as the frequency increases. Yet the static  $\alpha_{yy} = 9.978$  a.u. reported by Osted *et al.* is larger than the experimental value of  $\alpha_{yy} = 9.55$  a.u. at  $\omega = 0.088$  a.u.<sup>80</sup>, which cannot be the case. Lastly,  $\alpha_{xx}$  reported by Osted *et al.* appears to be underestimated, when compared to our results and the literature. We used the same geometry of the water molecule as Joshi *et al.* and Osted *et al.*, henceforth ruling it out as a possible source of discrepancies. In summary, water polarizabilities reported by Joshi *et al.* and Osted *et al.* are unreliable.

Table 5: Ground-state energy and electronic properties of water. All entries are in atomic units (a.u.).

Property	PSQMC (QZ4P)	RQMC (TZP) <sup>a</sup>	Literature
$E_0$	-76.426(3)	-76.342(2)	-76.4240(3) <sup>b</sup> , -76.4236(2) <sup>c</sup> , -76.4368(4) <sup>d</sup> , -76.437(3) <sup>e</sup> , -76.4388 <sup>f</sup>
$\mu_z$	0.748(2)	0.758(4)	0.7238 <sup>g</sup> , 0.7268 <sup>h</sup> , 0.7306 <sup>i</sup>
$\Theta_{xx}$	1.906(2)	1.918(8)	1.912 <sup>g</sup> , 1.96(2) <sup>j</sup>
$\Theta_{yy}$	-1.800(5)	-1.878(7)	-1.804 <sup>g</sup> , -1.86(2) <sup>j</sup>
$\Theta_{zz}$	-0.109(4)	-0.091(6)	-0.108 <sup>g</sup> , -0.10(3) <sup>j</sup>
$\Omega_{xxz}$	3.21(2)	3.62(1)	3.25 <sup>g</sup> , 3.206 <sup>k</sup>
$\Omega_{yyz}$	-1.32(2)	-1.51(2)	-1.34 <sup>g</sup> , -1.332 <sup>k</sup>
$\Omega_{zzz}$	-1.89(3)	-2.13(2)	-1.91 <sup>g</sup> , -1.874 <sup>k</sup>
$\langle 1/r_O \rangle_e$	23.462(6)	23.27(1)	23.46 <sup>k</sup>
$\langle 1/r_H \rangle_e$	5.7902(8)	5.763(2)	5.773 <sup>k</sup>
$\langle r^2 \rangle_e$	19.40(1)	19.653(6)	19.73 <sup>k</sup> , 18(2) <sup>l</sup>
$\langle qx/r_O^3 \rangle$	-0.024(8)	—	0.0
$\langle qy/r_O^3 \rangle$	-0.019(4)	—	0.0
$\langle qz/r_O^3 \rangle$	0.08(2)	-0.01(1)	0.0
$\langle qx/r_H^3 \rangle$	0.000(1)	—	0.0
$\langle qy/r_H^3 \rangle$	-0.0030(9)	—	0.0
$\langle qz/r_H^3 \rangle$	0.0010(8)	—	0.0
$q_O(xx)$	1.44(5)	1.86(5)	1.59 <sup>m</sup> , 1.455(5) <sup>n</sup>
$q_O(yy)$	-1.61(3)	-1.8(1)	-1.79 <sup>m</sup> , -1.66(1) <sup>n</sup>
$q_O(zz)$	0.20(5)	0.35(6)	0.20 <sup>m</sup> , 0.21(1) <sup>n</sup>
$q_H(yy)$	-0.262(9)	-0.30(2)	-0.2932 <sup>m</sup> , -0.2588(3) <sup>n</sup>
$q_H(aa)$	0.482(6)	0.52(2)	0.5175 <sup>m</sup> , 0.4559(2) <sup>n</sup>
$q_H(bb)$	-0.198(5)	-0.22(2)	-0.2243 <sup>m</sup> , -0.1971(2) <sup>n</sup>
$\alpha(^{\circ})$	1.16(2)	0.93(2)	1.031 <sup>m</sup> , 1.27 <sup>n</sup>
$\alpha_{xx}$	9.7(3)	—	9.944 <sup>o</sup> , 9.88 <sup>p</sup> , 9.74 <sup>q</sup> , 9.18 <sup>r</sup>
$\alpha_{yy}$	8.8(3)	—	9.052 <sup>o</sup> , 8.92 <sup>p</sup> , 9.15 <sup>q</sup> , 7.90 <sup>r</sup>
$\alpha_{zz}$	9.4(1)	—	9.398 <sup>o</sup> , 9.52 <sup>p</sup> , 9.36 <sup>q</sup> , 8.52 <sup>r</sup>

<sup>a</sup> Reptation quantum Monte Carlo, Reference 83, Table 3.<sup>b</sup> Diffusion quantum Monte Carlo, Reference 62.<sup>c</sup> Single-determinant diffusion Monte Carlo, Reference 76.<sup>d</sup> Multi-determinant diffusion Monte Carlo, Reference 76.<sup>e</sup> Explicitly correlated coupled cluster (CCSDTQ-R12), Reference 84.<sup>f</sup> Experiment, Reference 85.<sup>g</sup> SCF with electron correlation correction from CCSD(T), Reference 82.<sup>h</sup> Experiment, Reference 86.<sup>i</sup> Experiment, Reference 87.<sup>j</sup> Experiment, Reference 88.<sup>k</sup> CBS FCI, Reference 89. The value for  $\Omega_{xxz}$  is calculated from  $\Omega_{yyz}$  and  $\Omega_{zzz}$ .<sup>l</sup> Experiment, Reference 90.<sup>m</sup> CISD, Reference 91.<sup>n</sup> Experiment, Reference 92.<sup>o</sup> CCSD, Reference 79.<sup>p</sup> DFT, Reference 80.<sup>q</sup> MC SCF, Reference 81.<sup>r</sup> SCF, Reference 82.

### 4.3 Application to Ethylene

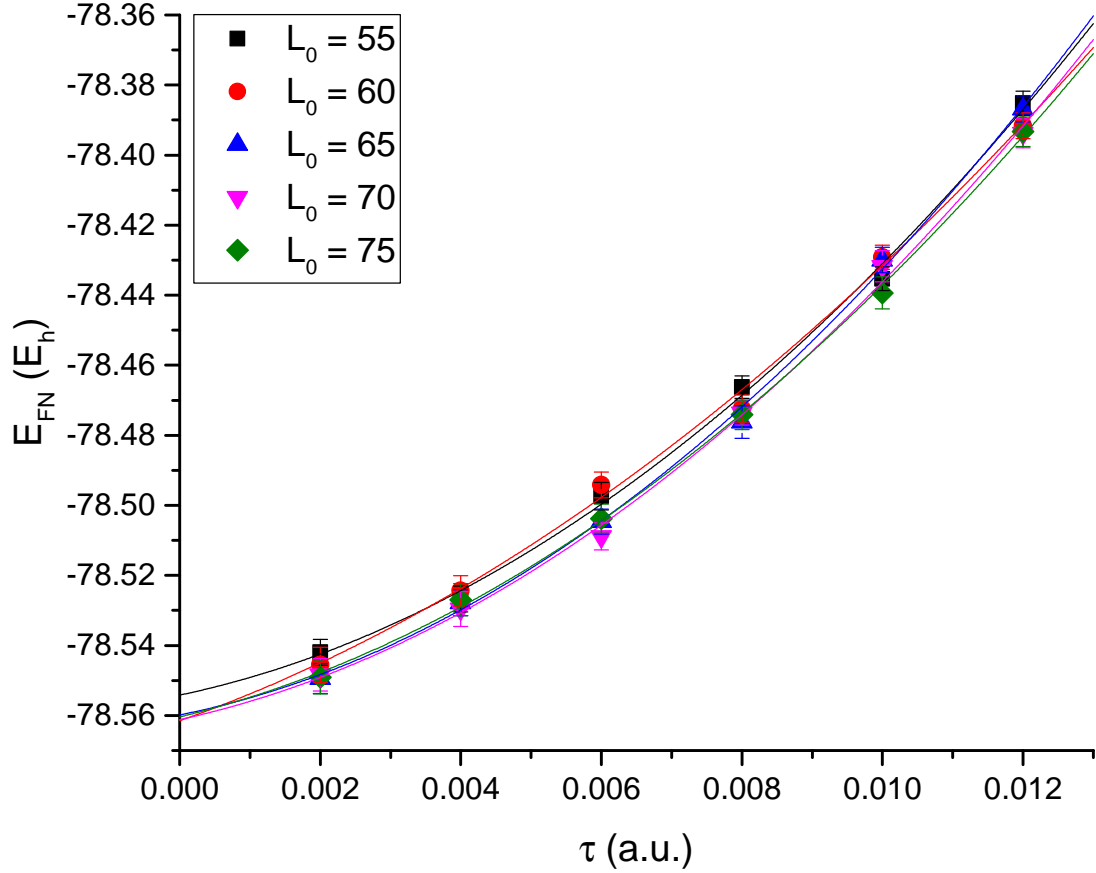


Figure 9: Ground-state energy of ethylene as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (46), where  $\tau_0 = 0.012$  a.u. Convergence occurs for  $L_0 = 60 \dots 75$ [5].

Our calculations for ethylene used experimental geometry from Reference 93. For each path-length parameter, we performed 100 independent runs with 20 000 iterations per run and with an additional 5 000 iterations thrown out for the purpose of equilibration. Table 7 contains our results for ethylene. The ground-state energy was calculated by variance-weighted averaging of converged results from four separate values of the path-length parameter,  $L_0 = 60 \dots 75$ [5]; see Figure 9 and Table 6 for details. The ground-state energy is in good agreement with Nemec *et al.*,<sup>62</sup> who used an identical trial wave function. Our result is above the exact ground-state energy, as expected from our obedience of the variational upper bound, but we still manage to

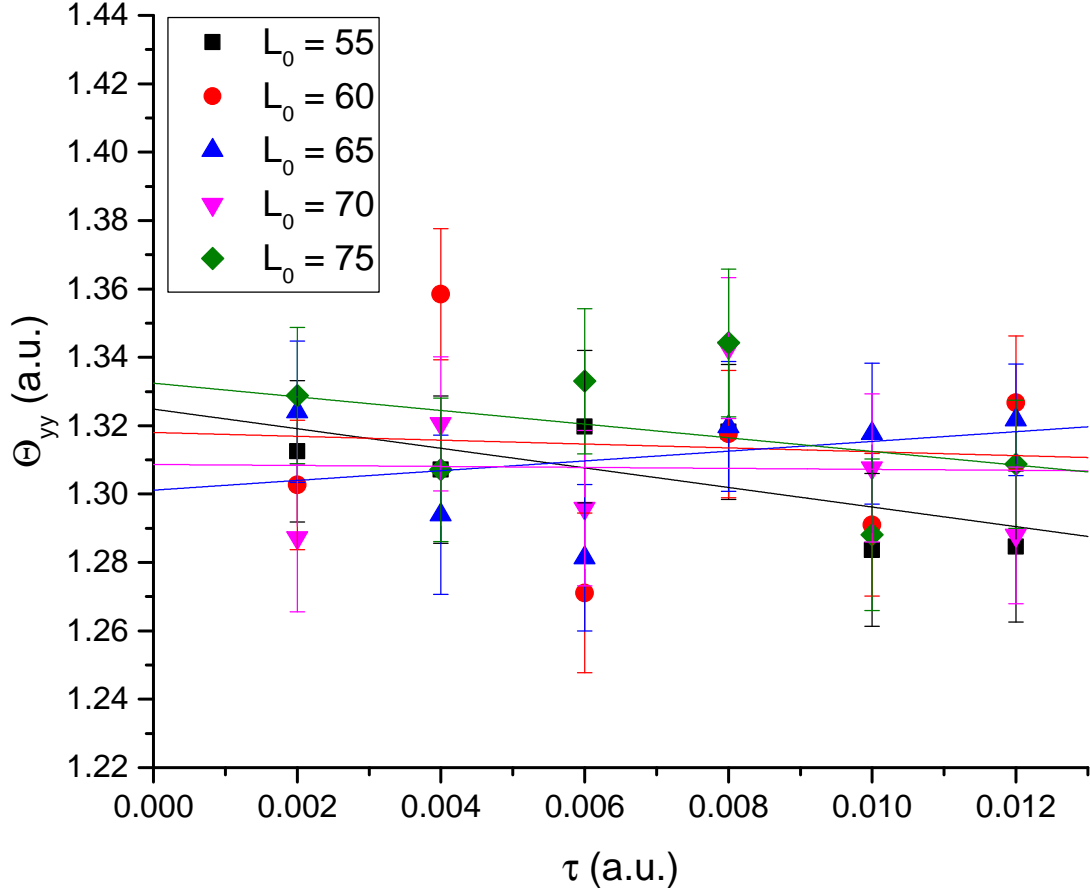


Figure 10: Quadrupole moment of ethylene as a function of the initial path-length parameter ( $L_0$ ). Path-lengths increase as time-steps decrease in accordance with Equation (46), where  $\tau_0 = 0.012$  a.u. Convergence occurs for  $L_0 = 60 \dots 75$ [5].

recover more electron correlation energy than the CCSD(T) literature, which is not variationally bound.

We illustrate the convergence of electronic properties using the yy-component of the quadrupole moment, which was obtained by variance-weighted averaging of converged results from four distinct values of the path-length parameter,  $L_0 = 60 \dots 75$ [5]; see Figure 10 and Table 6 for details. The figures illustrating the convergence of the remaining electronic properties of ethylene are shown in Appendix 5. It is difficult to say how accurate our quadrupole moment is, since the literature presents a diverse set of results and the theoretical methods employed different geometries in their calculations. Nonetheless, our results are consistent with symmetry restrictions and do not

Table 6: Ground-state energy and yy-component of the quadrupole moment of ethylene in the limit of zero time-step ( $\tau$ ) and as a function of path-length parameter ( $L_0$ ). All entries are in atomic units (a.u.).

$L_0$	$E_0$	$\Theta_{yy}$
55	$-78.5542(67)^\dagger$	$1.325(12)^\ddagger$
60	$-78.5615(87)$	$1.318(30)$
65	$-78.5598(57)$	$1.301(17)$
70	$-78.5612(66)$	$1.309(23)$
75	$-78.5604(38)$	$1.332(19)$
Average*	$-78.5605(27)$	$1.314(10)$

<sup>†</sup> Not included in the weighted average due to a lack of convergence with the rest of the data.

<sup>‡</sup> Not included in the weighted average because the ground-state energy at this path-length did not converge.

\* Variance-weighted average of the property for all included path-lengths. These properties and their corresponding literature values are also quoted in Table 5.

show any irregularities. The electric fields agree with their theoretical values, except for the  $x$  and  $y$  components on the hydrogen nuclei, which may reflect our use of an experimental geometry, instead of one that minimizes the Hartree-Fock energy of the importance sampling function. To our knowledge, diamagnetic shielding, diamagnetic susceptibility and electric field gradients of ethylene have not been calculated previously. We were not yet successful in calculating the static  $\alpha$  polarizability of ethylene and it remains a target of our future work.

Table 7: Ground-state energy and electronic properties of ethylene. All entries are in atomic units (a.u.).

Property	PSQMC	Literature
$E_0$	-78.561(3)	-78.567(0.4) <sup>a</sup> , -78.5526 <sup>b</sup> , -78.5888 <sup>c</sup>
$\Theta_{xx}$	-2.81(1)	-2.444 <sup>d</sup> , -2.5988 <sup>e</sup> , -2.8419 <sup>f</sup> -2.35(14) <sup>g</sup> , -2.45 <sup>h</sup> , -2.68 <sup>i</sup>
$\Theta_{yy}$	1.31(1)	1.196 <sup>d</sup> , 1.2332 <sup>e</sup> , 1.2399 <sup>f</sup> 1.10(7) <sup>g</sup> , 0.97 <sup>h</sup> , 1.04 <sup>i</sup>
$\Theta_{zz}$	1.500(8)	1.247 <sup>d</sup> , 1.3656 <sup>e</sup> , 1.6020 <sup>f</sup> 1.24(14) <sup>g</sup> , 1.49 <sup>h</sup> , 1.63 <sup>i</sup>
$\langle 1/r_C \rangle_e$	18.484(7)	—
$\langle 1/r_H \rangle_e$	6.2049(4)	—
$\langle r^2 \rangle_e$	83.23(3)	—
$\langle qx/r_C^3 \rangle$	-0.002(1)	0.0
$\langle qy/r_C^3 \rangle$	-0.002(1)	0.0
$\langle qz/r_C^3 \rangle$	-0.002(1)	0.0
$\langle qx/r_H^3 \rangle$	-0.0021(3)	0.0
$\langle qy/r_H^3 \rangle$	-0.0039(9)	0.0
$\langle qz/r_H^3 \rangle$	-0.0003(7)	0.0
$q_C(xx)$	0.297(6)	—
$q_C(yy)$	-0.094(7)	—
$q_C(zz)$	-0.206(7)	—
$q_H(xx)$	-0.164(2)	—
$q_H(yy)$	0.305(1)	—
$q_H(zz)$	-0.142(1)	—

The tensor  $q_H$  was diagonalized.

<sup>a</sup> Diffusion quantum Monte Carlo, Reference 62.

<sup>b</sup> CCSD(T), Reference 94.

<sup>c</sup> Exact non-relativistic electronic energy, Reference 95.

<sup>d</sup> CC3, Reference 96.

<sup>e</sup> MP2, Reference 97.

<sup>f</sup> SCF, Reference 97.

<sup>g</sup> Experiment, Reference 98.

<sup>h</sup> Experiment, Reference 99.

<sup>i</sup> Experiment, Reference 100.



#### 4.4 Comparison with RQMC

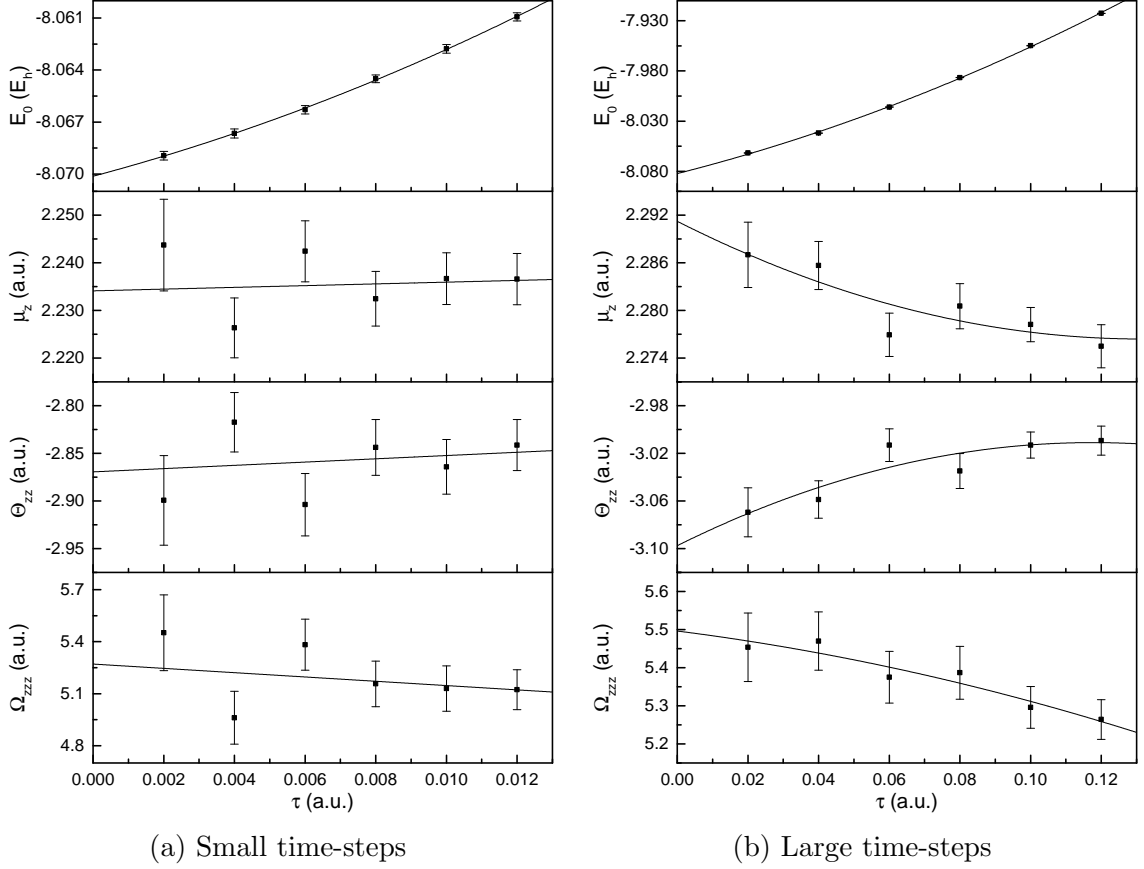


Figure 11: Properties of lithium hydride calculated using RQMC with reptile-length parameter  $L_0 = 121$ . Reptile-lengths increase as time-steps decrease in accordance with Equation (45). (a) The energy is being accurately sampled at the chosen set of time-steps, but the other properties are not;  $E_0 = -8.0701(4)$ ,  $\mu_z = 2.234(7)$ ,  $\Theta_{zz} = -2.87(4)$ ,  $\Omega_{zzz} = 5.27(18)$ . (b) The energy value is nonsensical and disobeys the variational upper bound, although the other properties are being accurately sampled at the chosen set of time-steps;  $E_0 = -8.082(3)$ ,  $\mu_z = 2.291(6)$ ,  $\Theta_{zz} = -3.10(3)$ ,  $\Omega_{zzz} = 5.50(14)$ .

We performed 100 independent runs on lithium hydride using reptation quantum Monte Carlo (RQMC) with algorithmic parameters reported in Reference 57: path-length (*total* length of the reptile) at the largest time-step ( $L_0$ ) equal to 121, chop-size at the largest time-step ( $M_0$ ) equal to 20, and 400 000 iterations per run. The reptile-length  $L$  at a given time-step is related to  $L_0$  by Equation (45). For the set of time-steps shown in Figure 11a, the energy is being accurately sampled, but the remaining

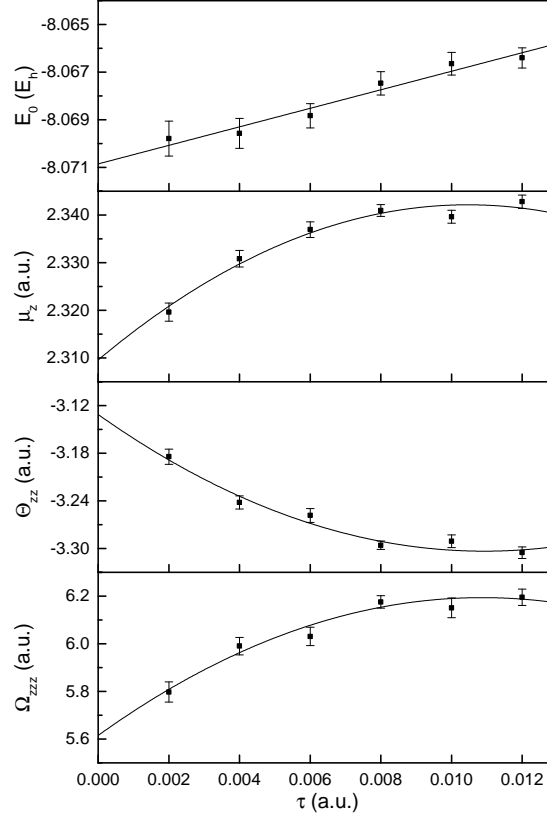


Figure 12: Properties of lithium hydride calculated using PSQMC with path-length parameter  $L_0 = 201$ . The total length of the sampled path, either  $X$  or  $Y$ , at the largest time-step is 403. Path-lengths increase as time-steps decrease in accordance with Equation (45). Here, the energy and other properties are accurately sampled *simultaneously* at the chosen set of time-steps;  $E_0 = -8.0709(4)$ ,  $\mu_z = 2.310(4)$ ,  $\Theta_{zz} = -3.13(2)$ ,  $\Omega_{zzz} = 5.62(8)$ .

properties are not. The reverse is true for the time-steps shown in Figure 11b. The other properties are being accurately sampled, but there is too much time-step bias for the energy to be accurately determined. Its value in the limit of zero time-step violates the variational energy upper bound.

Another set of 100 independent runs on lithium hydride was performed using pure-sampling quantum Monte Carlo (PSQMC), where  $L_0 = 201$  and each run consumed 20 000 iterations. The results are shown in Figure 12. Here,  $L_0$  is the length of paths  $X$  and  $Y$ , emanating from the middle to their ends, and the total length of path  $Z$ , each at the largest time-step; see Figure 1. Again, the path-length  $L$  at a given time-

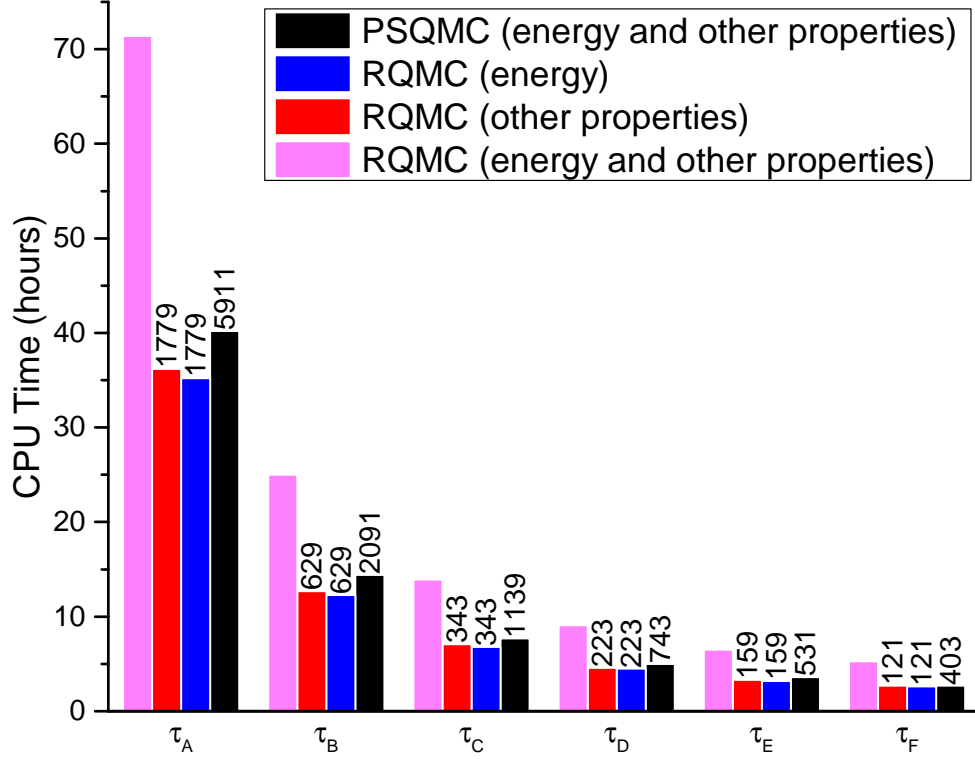


Figure 13: CPU time versus time-step for RQMC and PSQMC runs shown in Figures 11 and 12. RQMC reptile-lengths and PSQMC path-lengths at corresponding values of the time-steps are shown above the bars.  $\tau_A \dots \tau_F$  are 0.002...0.012[0.002] a.u. for PSQMC (energy and other properties) and RQMC (energy).  $\tau_A \dots \tau_F$  are 0.02...0.12[0.02] a.u. for RQMC (other properties).

step is related to  $L_0$  by Equation (45). In this case, the energy and other properties are accurately sampled *simultaneously* at the chosen set of time-steps. (We alert the reader to the fact that property values quoted in Figure 12 are based on runs with a single choice of  $L_0$ , while those reported in Table 2 are a variance-weighted average over multiple values of  $L_0$ ).

In Figure 13, we display the CPU time per run for RQMC and PSQMC, both performed on an Intel Xeon 2.2 GHz CPU at each time-step, for the cases where accurate sampling is observed. PSQMC is seen to be more computationally-efficient than RQMC by nearly a factor of two. In order to achieve the same efficiency as PSQMC, one could reduce the number of RQMC iterations. However, this would increase the RQMC error bars, which are already larger than those of PSQMC.

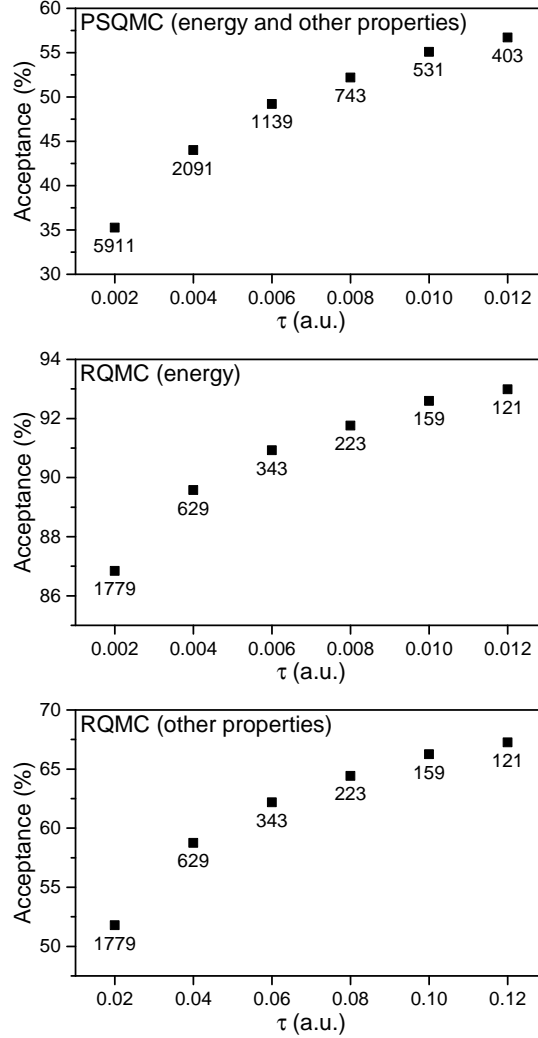


Figure 14: Metropolis acceptance rates versus time-steps for RQMC and PSQMC, corresponding to runs shown in Figures 11 and 12. RQMC reptile-lengths and PSQMC path-lengths at corresponding values of the time-steps are indicated below the points.

The method used to construct the paths in PSQMC has the beneficial effect of reducing the serial correlation for pure-sampling, but at the cost of reduced acceptance rates relative to those of RQMC. A comparison is made for lithium hydride in Figure 14. As expected, the acceptance rates decrease as the path-lengths increase, but the acceptance rate for PSQMC at the smallest time-step (largest path-length) is still large enough to sample configuration space effectively. The same holds true for water: PSQMC acceptance rates at the smallest time-step range from 58% to 48% for runs with  $L_0 = 201$  to 301, respectively.

## 5 Conclusion

We described a pure-sampling quantum Monte Carlo (PSQMC) algorithm that accurately samples properties from both mixed and pure electron distributions *simultaneously* and maintains the simplicity of fixed-node diffusion Monte Carlo, which samples only from the mixed distribution. In common with reptation quantum Monte Carlo (RQMC), the heart of our new algorithm is the Metropolis decision. This ensures that PSQMC is free of population control bias and provides stability when it is employed with large path-lengths. Conversely, PSQMC does not use reptation to generate new paths. Our implementation of the new algorithm converges to a threshold value of the path-length parameter. Averaging results taken at the threshold value and beyond provides accurate estimates of properties that are free from time-step and regression-model bias. In addition, the observed convergence gives us confidence in our choice of algorithmic parameters. The only source of error in our algorithm is the fixed-node approximation, which stems from the mismatch of the nodal hypersurface of the importance sampling function with that of the exact ground-state wave function.

For the purpose of sampling both mixed and pure distributions *simultaneously* in runs performed at a single set of time-steps, PSQMC is more efficient than RQMC, but one expects the Metropolis acceptance rates for PSQMC to be less than those for RQMC. For small molecules considered here, the Metropolis acceptance rates are still sufficiently large to properly sample configuration space. Work is in progress in our laboratory to explore this issue in applications to large molecules.

In our application of PSQMC, we found excellent agreement with the accepted values for the energy and a variety of other properties for lithium hydride, water and ethylene molecules. For lithium hydride, we observed excellent agreement between the experimental results and our values for the dipole moment and the electric field gradients. In the case of water, our dipole moment overestimates the experiment by about 10% but the electric field gradients are in good agreement with the experi-

mental results. It is difficult to compare our quadrupole moment of ethylene with experimental literature, due to conflicting results, but it does agree with at least one of them. We were able to provide insight in the disagreement between various literature results for the static  $\alpha$  polarizability of water. Finally, we reported several properties of ethylene that have not been previously reported in the literature.

Current and future work in our lab involves performing PSQMC calculations on large molecules, such as polyacetylene and polyfuran, in combination with importance sampling functions obtained from Density Functional Theory. We wish to use the strengths of our algorithm, namely unambiguous convergence in path-length and its ability to perform pure-sampling, to investigate the nodal hypersurfaces of these functions and identify the ones that have the best nodes.

Other methods that we foresee enjoying more attention and development in the future are path integral ground state (PIGS)<sup>43–45</sup> and full configuration interaction quantum Monte Carlo (FCIQMC)<sup>33,34</sup>, both of which are capable of pure-sampling. The latter method is especially worthwhile, because it does not rest upon the fixed-node approximation and it is capable of calculating excited states.<sup>33,34</sup>

Another direction that should be explored is the use of graphics processing units (GPUs) to accelerate quantum Monte Carlo (QMC) calculations. Some groups have already started to utilize GPUs in their QMC calculations<sup>101–104</sup>, reporting a substantial performance increase of up to 150 times.<sup>104</sup> Nonetheless, there is plenty of room for further performance increases. This can be evident from the field of molecular dynamics, which started to experiment with GPUs much earlier and achieved a performance increase in excess of 700 times.<sup>105</sup> Ultimately, using GPUs will increase the efficiency of QMC methods and allow us to study much larger and more challenging systems.

## 6 References

- [1] Hertz, H. *Ann. Phys.* **267**, 983–1000 (1887).
- [2] Einstein, A. *Ann. Phys.* **322**, 132–148 (1905).
- [3] Compton, A. H. *Phys. Rev.* **21**, 483–502 (1923).
- [4] Metropolis, N. and Ulam, S. *J. Am. Stat. Assoc.* **44**, 335–341 (1949).
- [5] Kalos, M. H. *Phys. Rev.* **128**, 1791–1795 (1962).
- [6] Anderson, J. B. *J. Chem. Phys.* **63**, 1499–1503 (1975).
- [7] Rothstein, S. M. *Can. J. Chem.* **91**, 505–510 (2013).
- [8] Austin, B. M., Zubarev, D. Y., and Lester Jr., W. A. *Chem. Rev.* **112**, 263–288 (2012).
- [9] Foulkes, W. M. C., Mitas, L., Needs, R. J., and Rajagopal, G. *Rev. Mod. Phys.* **73**, 33–83 (2001).
- [10] Ceperley, D. M. and Bernu, B. *J. Chem. Phys.* **89**, 6316–6328 (1988).
- [11] Bernu, B., Ceperley, D. M., and Lester, W. A. *J. Chem. Phys.* **93**, 552–561 (1990).
- [12] Umrigar, C. J., Nightingale, M. P., and Runge, K. J. *J. Chem. Phys.* **99**, 2865–2890 (1993).
- [13] Baroni, S. and Moroni, S. *Phys. Rev. Lett.* **82**, 4745–4748 (1999).
- [14] Anderson, J. B. *J. Chem. Phys.* **65**, 4121–4127 (1976).
- [15] Yuen, W. K., Farrar, T. J., and Rothstein, S. M. *J. Phys. A* **40**, F639–F646 (2007).

- 
- [16] Hohenberg, P. and Kohn, W. *Phys. Rev.* **136**, B864–B871 (1964).
- [17] Kohn, W. and Sham, L. J. *Phys. Rev.* **140**, A1133–A1138 (1965).
- [18] Kohn, W., Becke, A. D., and Parr, R. G. *J. Phys. Chem.* **100**, 12974–12980 (1996).
- [19] Becke, A. D. *J. Chem. Phys.* **140**, 18A301 (2014).
- [20] Coester, F. and Kümmel, H. *Nucl. Phys.* **17**, 477–485 (1960).
- [21] Čížek, J. *J. Chem. Phys.* **45**, 4256–4266 (1966).
- [22] Paldus, J., Čížek, J., and Shavitt, I. *Phys. Rev. A* **5**, 50–67 (1972).
- [23] Purvis III, G. D. and Bartlett, R. J. *J. Chem. Phys.* **76**, 1910–1918 (1982).
- [24] Bartlett, R. J. and Musiał, M. *Rev. Mod. Phys.* **79**, 291–352 (2007).
- [25] Bacher, R. F. *Phys. Rev.* **43**, 264–269 (1933).
- [26] Ufford, C. W. *Phys. Rev.* **44**, 732–739 (1933).
- [27] Shavitt, I. *Mol. Phys.* **94**, 3–17 (1998).
- [28] Sherrill, C. D. and Schaefer III, H. F. *Adv. Quantum Chem.* **34**, 143–269 (1999).
- [29] Szalay, P. G., Müller, T., Gidofalvi, G., Lischka, H., and Shepard, R. *Chem. Rev.* **112**, 108–181 (2012).
- [30] Anderson, J. B. *Int. J. Quantum Chem.* **15**, 109–120 (1979).
- [31] Klein, D. J. and Pickett, H. M. *J. Chem. Phys.* **64**, 4811–4812 (1976).
- [32] Ceperley, D. M. *J. Stat. Phys.* **63**, 1237–1267 (1991).
- [33] Booth, G. H., Thom, A. J. W., and Alavi, A. *J. Chem. Phys.* **131**, 054106 (2009).



- 
- [34] Blunt, N. S., Smart, S. D., Booth, G. H., and Alavi, A. *J. Chem. Phys.* **143**, 134117 (2015).
- [35] Ospadov, E. and Rothstein, S. M. *J. Chem. Phys.* **142**, 024114 (2015).
- [36] Born, M. and Oppenheimer, R. *Ann. Phys.* **389**, 457–484 (1927).
- [37] Trotter, H. F. *Proc. Am. Math. Soc.* **10**, 545 (1959).
- [38] Suzuki, M. *Commun. Math. Phys.* **51**, 183–190 (1976).
- [39] Reynolds, P. J., Ceperley, D. M., Alder, B. J., and Lester Jr., W. A. *J. Chem. Phys.* **77**, 5593–5603 (1982).
- [40] Rothstein, S. M. and Vrbik, J. *J. Chem. Phys.* **87**, 1902–1903 (1987).
- [41] Vrbik, J. and Rothstein, S. M. *J. Comput. Phys.* **63**, 130–139 (1986).
- [42] Rothstein, S. M., Patil, N., and Vrbik, J. *J. Comput. Chem.* **8**, 412–419 (1987).
- [43] Sarsa, A., Schmidt, K. E., and Magro, W. R. *J. Chem. Phys.* **113**(4), 1366–1371 (2000).
- [44] Cuervo, J. E., Roy, P. N., and Boninsegni, M. *J. Chem. Phys.* **122**, 114504 (2005).
- [45] Cuervo, J. E. and Roy, P. N. *J. Chem. Phys.* **125**, 124314 (2006).
- [46] Liu, K. S., Kalos, M. H., and Chester, G. V. *Phys. Rev. A* **10**, 303–308 (1974).
- [47] Ceperley, D. M. and Alder, B. J. *Phys. Rev. Lett.* **45**, 566–569 (1980).
- [48] Ceperley, D. M. and Alder, B. J. *J. Chem. Phys.* **81**, 5833–5844 (1984).
- [49] Chen, B. and Anderson, J. B. *J. Chem. Phys.* **102**, 4491–4494 (1995).
- [50] Anderson, J. B. *Int. Rev. Phys. Chem.* **14**, 85–112 (1995).

- [51] Jastrow, R. *Phys. Rev.* **98**, 1479–1484 (1955).
- [52] DePasquale, M. F., Rothstein, S. M., and Vrbik, J. *J. Chem. Phys.* **89**, 3629–3637 (1988).
- [53] Buckingham, A. D. *Quart. Rev., Chem. Soc.* **13**, 183–214 (1959).
- [54] Kielich, S. and Zawodny, R. *Chem. Phys. Lett.* **12**, 20–24 (1971).
- [55] Smith, J. A. S. *J. Chem. Educ.* **48**, 39–48 (1971).
- [56] Vrbik, J., Legare, D. A., and Rothstein, S. M. *J. Chem. Phys.* **92**, 1221–1227 (1990).
- [57] Ospadov, E., Oblinsky, D. G., and Rothstein, S. M. *Phys. Chem. Chem. Phys.* **13**, 8031–8036 (2011).
- [58] te Velde, G., Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G., and Ziegler, T. *J. Comput. Chem.* **22**, 931–967 (2001).
- [59] Fonseca Guerra, C., Snijders, J. G., te Velde, G., and Baerends, E. J. *Theor. Chem. Acc.* **99**, 391–403 (1998).
- [60] ADF2010, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, <http://www.scm.com>.
- [61] Herzberg, G. and Huber, K. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*. Van Nostrand, New York, (1979).
- [62] Nemec, N., Towler, M. D., and Needs, R. J. *J. Chem. Phys.* **132**, 034111 (2010).
- [63] East, A. L. L., Rothstein, S. M., and Vrbik, J. *J. Chem. Phys.* **89**, 4880–4884 (1988).

- [64] Yuen, W. K., Ospadov, E., and Rothstein, S. M. *Journal of Computational Science* **5**, 542–548 (2014).
- [65] Cencek, W. and Rychlewski, J. *Chem. Phys. Lett.* **320**, 549–552 (2000).
- [66] Tunega, D. and Noga, J. *Theor. Chem. Acc.* **100**, 78–84 (1998).
- [67] Cafiero, M. and Adamowicz, L. *Phys. Rev. Lett.* **88**, 033002 (2002).
- [68] Wharton, L., Gold, L. P., and Klemperer, W. *J. Chem. Phys.* **37**, 2149–2150 (1962).
- [69] Lee, B. K., Stout, J. M., and Dykstra, C. E. *J. Mol. Struct. THEOCHEM* **400**, 57–68 (1997).
- [70] Bishop, D. M. and Lam, B. *Chem. Phys. Lett.* **120**, 69–74 (1985).
- [71] Keil, F. and Ahlrichs, R. *J. Chem. Phys.* **71**, 2671–2675 (1979).
- [72] Arrighini, G., Tomasi, J., and Guidotti, C. *Theor. Chim. Acta* **18**, 329–340 (1970).
- [73] Karna, S. P. and Grein, F. *Mol. Phys.* **69**, 661–670 (1990).
- [74] Wharton, L., Gold, L. P., and Klemperer, W. *Phys. Rev.* **133**, B270–B272 (1964).
- [75] Herzberg, G. *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules*. Van Nostrand, New York, (1945).
- [76] Clark, B. K., Morales, M. A., McMinis, J., Kim, J., and Scuseria, G. E. *J. Chem. Phys.* **135**, 244105 (2011).
- [77] Joshi, S. P. and Vaval, N. *Chem. Phys. Lett.* **568–569**, 170–175 (2013).

- [78] Osted, A., Kongsted, J., Mikkelsen, K. V., and Christiansen, O. *J. Phys. Chem. A* **108**, 8646–8658 (2004).
- [79] Kongsted, J., Osted, A., Mikkelsen, K. V., and Christiansen, O. *J. Chem. Phys.* **118**, 1620–1633 (2003).
- [80] Jensen, L., van Duijnen, P. T., and Snijders, J. G. *J. Chem. Phys.* **119**, 3800–3809 (2003).
- [81] Poulsen, T. D., Ogilby, P. R., and Mikkelsen, K. V. *J. Chem. Phys.* **116**, 3730–3738 (2002).
- [82] Maroulis, G. *Chem. Phys. Lett.* **289**, 403–411 (1998).
- [83] Oblinsky, D. G., Yuen, W. K., and Rothstein, S. M. *J. Mol. Struct. THEOCHEM* **961**, 29–34 (2010).
- [84] Shiozaki, T., Kamiya, M., Hirata, S., and Valeev, E. F. *J. Chem. Phys.* **130**, 054101 (2009).
- [85] Klopper, W. *Mol. Phys.* **99**, 481–507 (2001).
- [86] Clough, S. A., Beers, Y., Klein, G. P., and Rothman, L. S. *J. Chem. Phys.* **59**, 2254–2259 (1973).
- [87] Shostak, S. L., Ebenstein, W. L., and Muentert, J. S. *J. Chem. Phys.* **94**, 5875–5882 (1991).
- [88] Verhoeven, J. and Dymanus, A. *J. Chem. Phys.* **52**, 3222–3233 (1970).
- [89] Feller, D. *J. Chem. Phys.* **98**, 7059–7071 (1993).
- [90] Eisenberg, D., Pochan, J. M., and Flygare, W. H. *J. Chem. Phys.* **43**, 4531–4532 (1965).

- [91] Rosenberg, B. J. and Shavitt, I. *J. Chem. Phys.* **63**, 2162–2174 (1975).
- [92] Verhoeven, J., Dymanus, A., and Bluysen, H. *J. Chem. Phys.* **50**, 3330–3338 (1969).
- [93] Herzberg, G. *Electronic Spectra and Electronic Structure of Polyatomic Molecules*. Van Nostrand, New York, (1966).
- [94] Machado, F. and Davidson, E. *J. Mol. Struct. THEOCHEM* **400**, 169–176 (1997).
- [95] O’Neill, D. and Gill, P. *Mol. Phys.* **103**, 763–766 (2005).
- [96] Hald, K., Halkier, A., Jørgensen, P., and Coriani, S. *J. Chem. Phys.* **117**, 9983–9990 (2002).
- [97] Russell, A. and Spackman, M. *Mol. Phys.* **98**, 855–865 (2000).
- [98] Majera, W., Lutzmann, P., and Hüttner, W. *Mol. Phys.* **83**, 567–578 (1994).
- [99] Bose, T. K., Boudjarane, K., Huot, J., and St-Arnaud, J. M. *J. Chem. Phys.* **89**, 7435–7440 (1988).
- [100] Kukolich, S. G., Aldrich, P. D., Read, W. G., and Campbell, E. J. *J. Chem. Phys.* **79**, 1105–1110 (1983).
- [101] Anderson, A. G., Goddard III, W. A., and Schröder, P. *Comput. Phys. Commun.* **177**, 298–306 (2007).
- [102] Uejima, Y., Terashima, T., and Maezono, R. *J. Comput. Chem.* **32**, 2264–2272 (2011).
- [103] Esler, K. P., Kim, J., Ceperley, D. M., and Shulenburger, L. *Comput. Sci. Eng.* **14**, 40–51 (2012).

- 
- [104] Lutsyshyn, Y. *Comput. Phys. Commun.* **187**, 162–174 (2015).
- [105] Friedrichs, M. S., Eastman, P., Vaidyanathan, V., Houston, M., Legrand, S., Beberg, A. L., Ensign, D. L., Bruns, C. M., and Pande, V. S. *J. Comput. Chem.* **30**, 864–872 (2009).

## A Appendix

### A.1 Derivation of the Modified Schrödinger Equation

To derive the modified Schrödinger equation presented in this work, let us start with the well-known, time-dependent Schrödinger equation:

$$\hat{H}\Phi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Phi(\mathbf{r}, t) \quad (47)$$

$$\left( -\frac{\hbar^2}{2m_e} \nabla^2 + V(\mathbf{r}) \right) \Phi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Phi(\mathbf{r}, t) \quad (48)$$

Next, rewrite this equation in imaginary time ( $it$ ) and atomic units. At the same time, introduce a constant shift in the zero-point energy, denoted by  $E_T$ . This will prove useful once we look for the ground-state solution to the modified Schrödinger equation. Accordingly, we obtain the following expression:

$$\left( -\frac{1}{2} \nabla^2 + V(\mathbf{r}) - E_T \right) \Phi(\mathbf{r}, t) = -\frac{\partial}{\partial t} \Phi(\mathbf{r}, t) \quad (49)$$

At this point, multiply both sides of this equation by the known guiding wave-function,  $\Psi(\mathbf{r})$ . Let us also suppress the dependence on  $\mathbf{r}$  and  $t$  for the sake of notational simplicity. By doing so, we arrive at:

$$\Psi \left( -\frac{1}{2} \nabla^2 + V - E_T \right) \Phi = -\Psi \frac{\partial \Phi}{\partial t} \quad (50)$$

$$\Psi \left( -\frac{1}{2} \nabla^2 + V - E_T \right) \Phi = -\frac{\partial}{\partial t} (\Phi \Psi) + \Phi \frac{\partial \Psi}{\partial t} \quad (51)$$

Since  $\Psi(\mathbf{r})$  is time-independent, then  $\Phi \frac{\partial \Psi}{\partial t} = 0$ . Using this result, we can write aforementioned equation as:

$$\Psi \left( -\frac{1}{2} \nabla^2 + V - E_T \right) \Phi = -\frac{\partial}{\partial t} (\Phi \Psi) \quad (52)$$

Using a vector identity,  $\nabla^2 (\Phi \Psi) = \Psi \nabla^2 \Phi + \Phi \nabla^2 \Psi + 2\nabla \Phi \cdot \nabla \Psi$ , express the previous

equation as:

$$-\frac{1}{2} (\nabla^2 (\Phi\Psi) - \Phi\nabla^2\Psi - 2\nabla\Phi \cdot \nabla\Psi) + V\Phi\Psi - E_T\Phi\Psi = -\frac{\partial}{\partial t} (\Phi\Psi) \quad (53)$$

$$-\frac{1}{2}\nabla^2 (\Phi\Psi) + \frac{1}{2}\Phi\nabla^2\Psi + \nabla\Phi \cdot \nabla\Psi + V\Phi\Psi - E_T\Phi\Psi = -\frac{\partial}{\partial t} (\Phi\Psi) \quad (54)$$

Applying yet another vector identity,  $\nabla \cdot (\Phi\nabla\Psi) = \nabla\Phi \cdot \nabla\Psi + \Phi\nabla^2\Psi$ , write it as:

$$-\frac{1}{2}\nabla^2 (\Phi\Psi) - \frac{1}{2}\Phi\nabla^2\Psi + \nabla \cdot (\Phi\nabla\Psi) + V\Phi\Psi - E_T\Phi\Psi = -\frac{\partial}{\partial t} (\Phi\Psi) \quad (55)$$

Finally, rearrange some terms in this equation to obtain:

$$-\frac{1}{2}\nabla^2 (\Phi\Psi) + \nabla \cdot \left( \Phi\Psi \frac{\nabla\Psi}{\Psi} \right) + \left( -\frac{1}{2} \frac{\nabla^2\Psi}{\Psi} + V - E_T \right) \Phi\Psi = -\frac{\partial}{\partial t} (\Phi\Psi) \quad (56)$$

By substituting the following three equations into the previous equation,

$$f(\mathbf{r}, t) = \Phi(\mathbf{r}, t)\Psi(\mathbf{r}) \quad (57)$$

$$\mathbf{F}(\mathbf{r}) = \frac{\nabla\Psi(\mathbf{r})}{\Psi(\mathbf{r})} \quad (58)$$

$$E_{loc}(\mathbf{r}) = -\frac{1}{2} \frac{\nabla^2\Psi(\mathbf{r})}{\Psi(\mathbf{r})} + V(\mathbf{r}) = \frac{\hat{H}\Psi(\mathbf{r})}{\Psi(\mathbf{r})} \quad (59)$$

we obtain the following expression:

$$-\frac{1}{2}\nabla^2 f(\mathbf{r}, t) + \nabla \cdot (f(\mathbf{r}, t)\mathbf{F}(\mathbf{r})) + (E_{loc}(\mathbf{r}) - E_T) f(\mathbf{r}, t) = -\frac{\partial f(\mathbf{r}, t)}{\partial t} \quad (60)$$

On the surface, this appears to be the modified Schrödinger equation, but it is not quite so. This equation still depends on time, unlike the equation presented in the Theory section. Furthermore, this equation depends on the arbitrary state  $\Phi(\mathbf{r}, t)$ , and not the ground-state  $\Phi_0$ . These discrepancies can be ameliorated by considering the general solution to the time-dependent Schrödinger equation in the form of a stationary state. Expressing it as a complete set of eigenfunctions  $\phi(\mathbf{r}, t)$



and corresponding energy eigenvalues  $E_i$  of the Hamiltonian, to obtain:

$$\Phi(\mathbf{r}, t) = \sum_{i=0}^{\infty} c_i e^{-(E_i - E_T)t} \phi_i(\mathbf{r}) \quad (61)$$

Only the term corresponding to the lowest energy state,  $E_0$ , will contribute to the sum in Equation (61) at sufficiently long times. Since we are trying to obtain the ground-state solution, the coefficient  $c_0 \neq 0$  and Equation (61) reduces to:

$$\Phi(\mathbf{r}, t) = c_0 e^{-(E_0 - E_T)t} \phi_0(\mathbf{r}) \quad (62)$$

At this point, we choose to fix  $E_T$  to be the true ground-state energy,  $E_T = E_0$ . In doing so, the aforementioned asymptotic solution is revealed to be a steady-state solution corresponding to the exact, time-independent, ground-state wave-function  $\phi_0(\mathbf{r})$ . We can define this solution as  $\Phi_0 = \phi_0(\mathbf{r})$  and use it in our definition for  $f(\mathbf{r}, t)$  to obtain:

$$f = f(\mathbf{r}) = \Phi_0(\mathbf{r})\Psi(\mathbf{r}) \quad (63)$$

Using this expression in conjunction with Equation (60), we finally recover the proposed, time-independent, modified Schrödinger equation:

$$-\frac{1}{2}\nabla^2 f + \nabla \cdot (f\mathbf{F}(\mathbf{r})) + (E_{loc}(\mathbf{r}) - E_0) f = -\frac{\partial f}{\partial t} = 0 \quad (64)$$

An important point to consider is the consequence of using the asymptotic limit with respect to time in order to obtain the modified Schrödinger equation and its solution presented above. Any method that relies on this equation must undergo sufficient equilibration in order to discard solutions corresponding to the mixture of excited and ground states, provided that one is trying to obtain the ground-state solution. Conversely, one could extract information about the excited states from such a mixture, which is a feat outside of the scope of this work.

## A.2 Derivation of the Energy Estimator

Here, we describe how to obtain Equation (25) and why sampling from the mixed distribution would yield the exact ground-state energy. Keep in mind that we cannot sample the energy from the pure distribution in our method, because we do not know  $\Phi_0$  exactly. We start by writing the time-independent Schrödinger equation:

$$\hat{H}|\Phi_0\rangle = E_0|\Phi_0\rangle \quad (65)$$

Next, we multiply both sides by  $\langle\Psi|$  and re-arrange for  $E_0$ :

$$\langle\Psi|\hat{H}|\Phi_0\rangle = \langle\Psi|E_0|\Phi_0\rangle \quad (66)$$

$$\langle\Psi|\hat{H}|\Phi_0\rangle = E_0\langle\Psi|\Phi_0\rangle \quad (67)$$

$$E_0 = \frac{\langle\Psi|\hat{H}|\Phi_0\rangle}{\langle\Psi|\Phi_0\rangle} \quad (68)$$

We can re-write this in integral representation, whilst supressing the complex conjugates since our wave functions are real.

$$E_0 = \frac{\int \Psi(\mathbf{r})\hat{H}\Phi_0(\mathbf{r})d\mathbf{r}}{\int \Psi(\mathbf{r})\Phi_0(\mathbf{r})d\mathbf{r}} = \frac{\int \Psi(\mathbf{r})\left(\frac{\hat{H}\Psi(\mathbf{r})}{\Psi(\mathbf{r})}\right)\Phi_0(\mathbf{r})d\mathbf{r}}{\int \Psi(\mathbf{r})\Phi_0(\mathbf{r})d\mathbf{r}} = \frac{\int \Psi(\mathbf{r})E_{loc}(\mathbf{r})\Phi_0(\mathbf{r})d\mathbf{r}}{\int \Psi(\mathbf{r})\Phi_0(\mathbf{r})d\mathbf{r}} \quad (69)$$

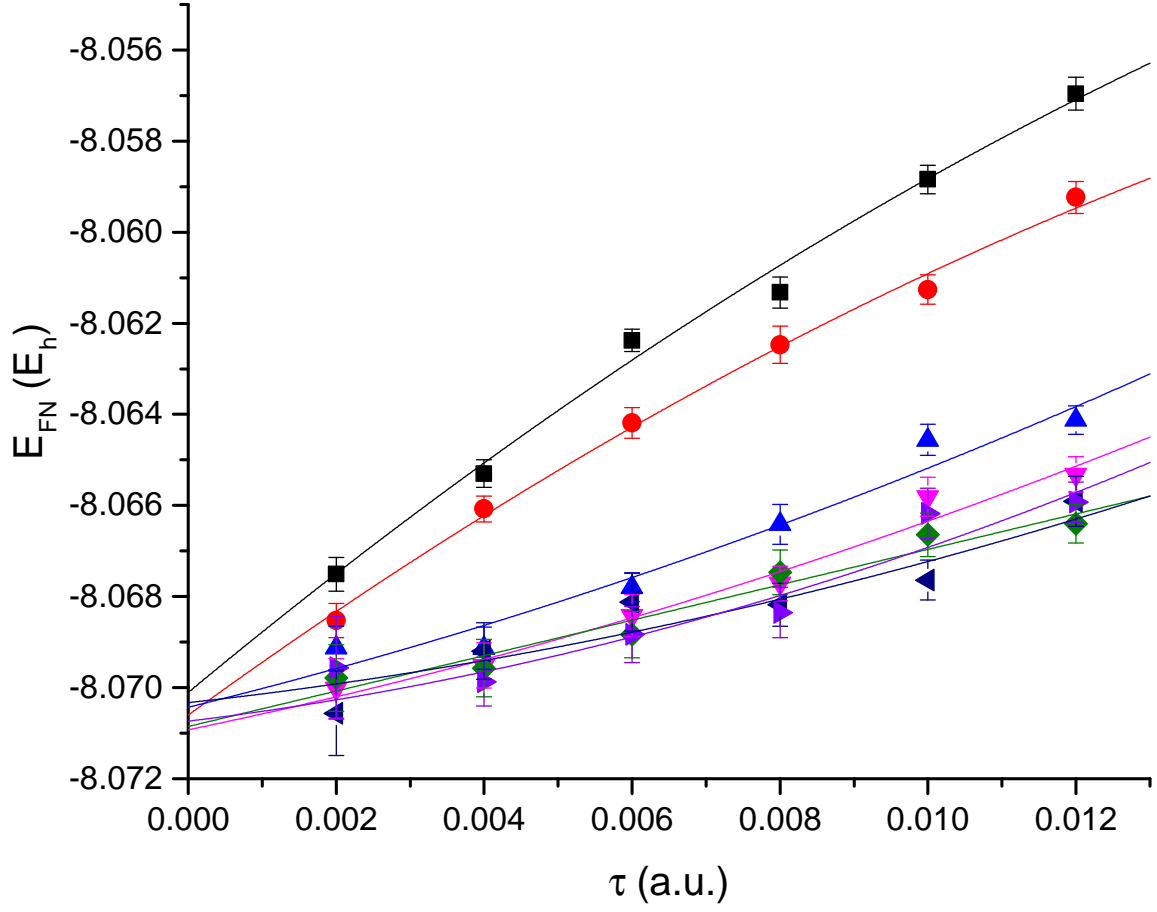
$$= \frac{\int \Psi(\mathbf{r})\Phi_0(\mathbf{r})E_{loc}(\mathbf{r})d\mathbf{r}}{\int \Psi(\mathbf{r})\Phi_0(\mathbf{r})d\mathbf{r}} = \frac{\int f(\mathbf{r})E_{loc}(\mathbf{r})d\mathbf{r}}{\int f(\mathbf{r})d\mathbf{r}} \quad (70)$$

where we used the fact that  $E_{loc}(\mathbf{r}) = \frac{\hat{H}\Psi(\mathbf{r})}{\Psi(\mathbf{r})}$  and  $f(\mathbf{r}) = \Psi(\mathbf{r})\Phi_0(\mathbf{r})$ . Finally, we can use the law of large numbers with the result from above to arrive at:

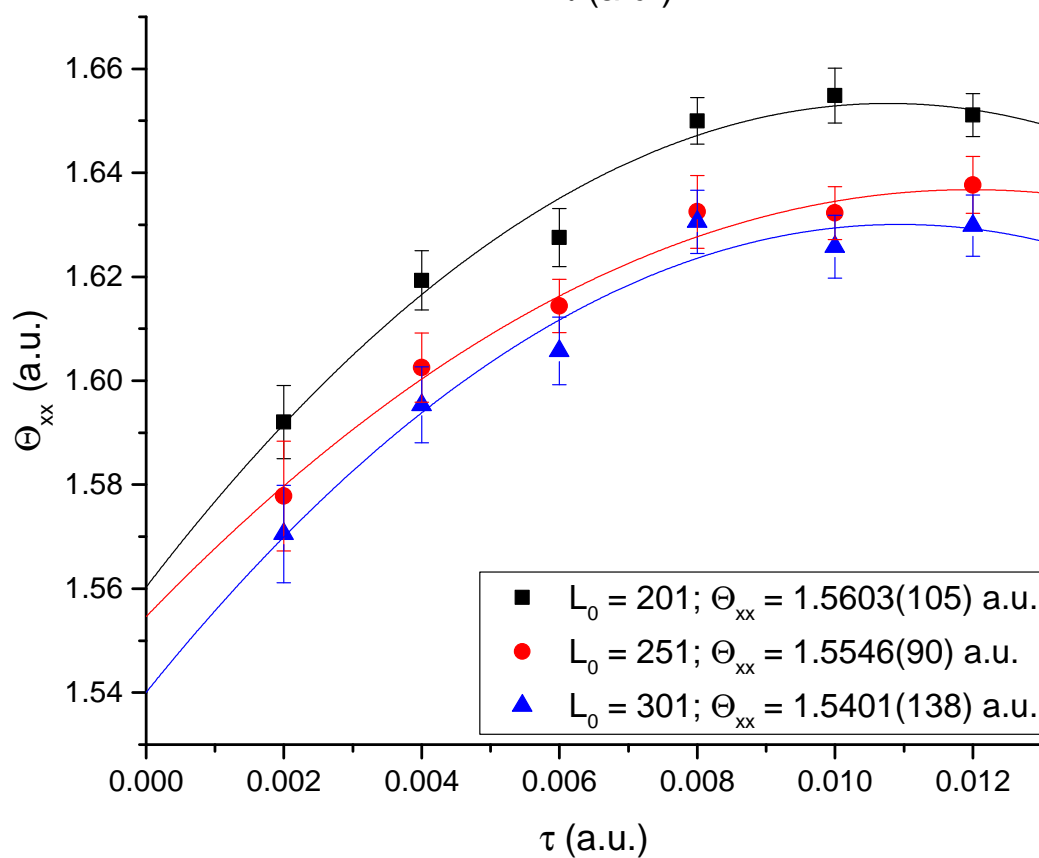
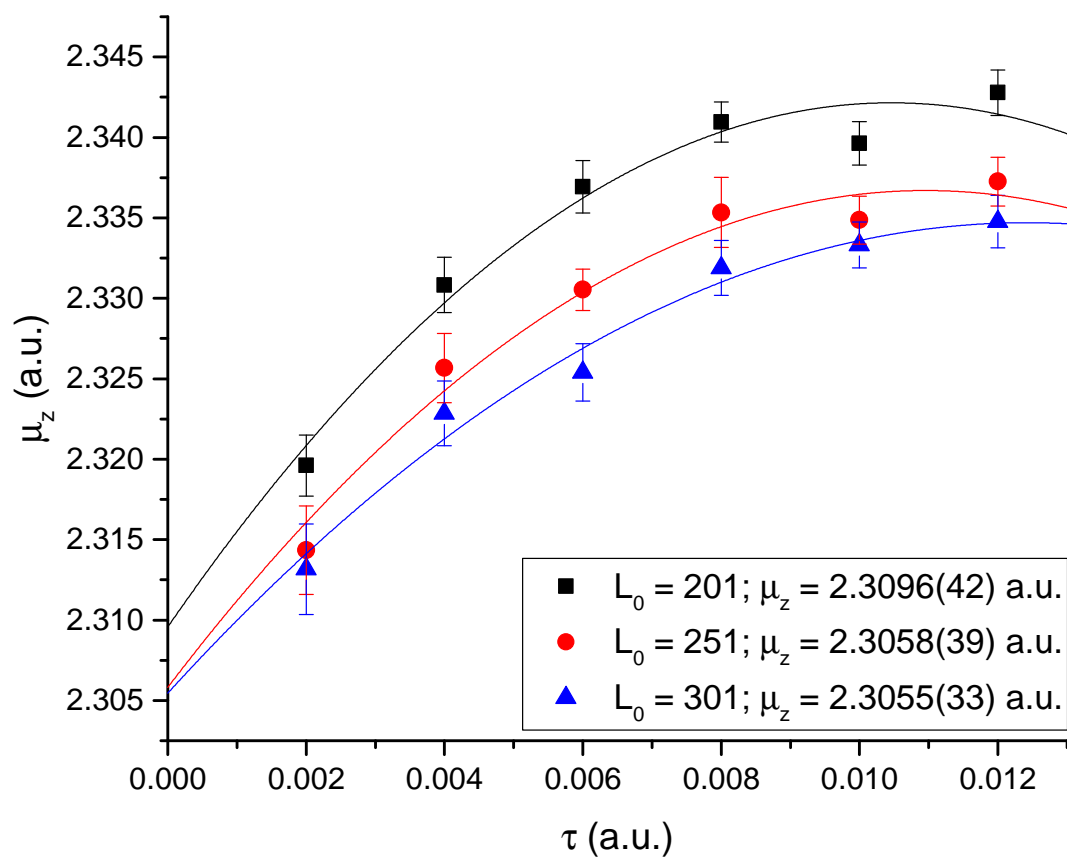
$$E_0 = \frac{1}{I} \sum_{i=1}^I E_{loc}(\mathbf{r}_i) \quad (71)$$

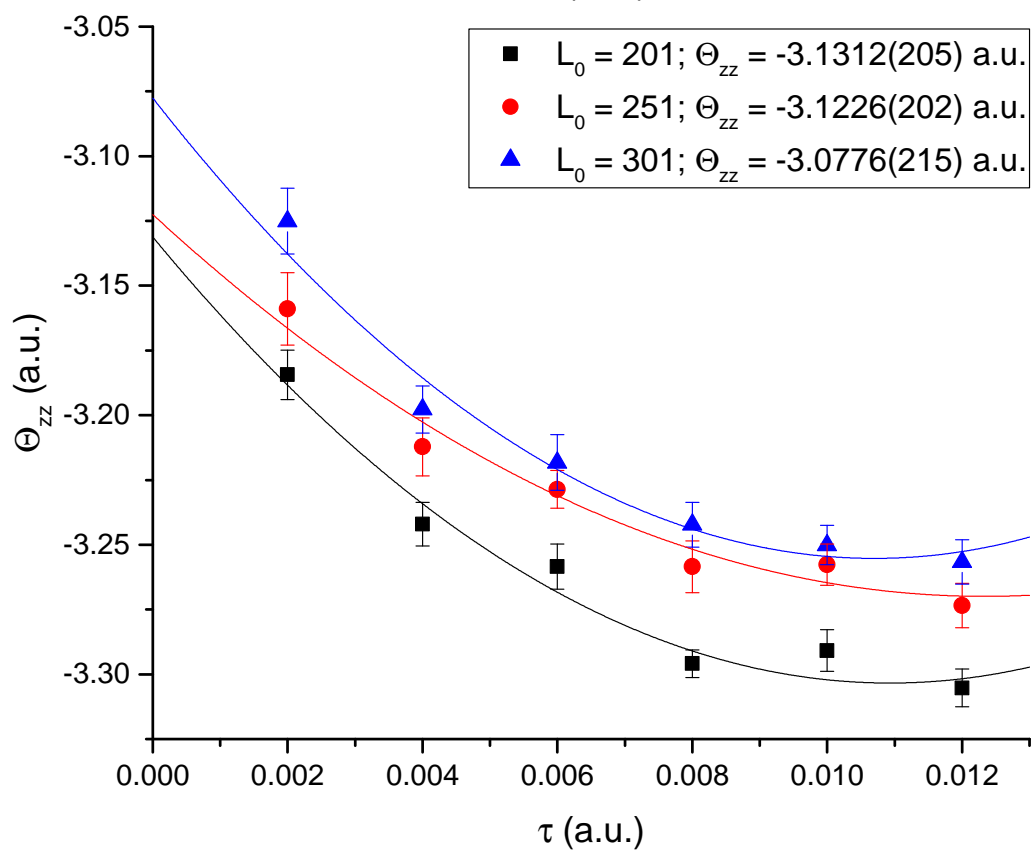
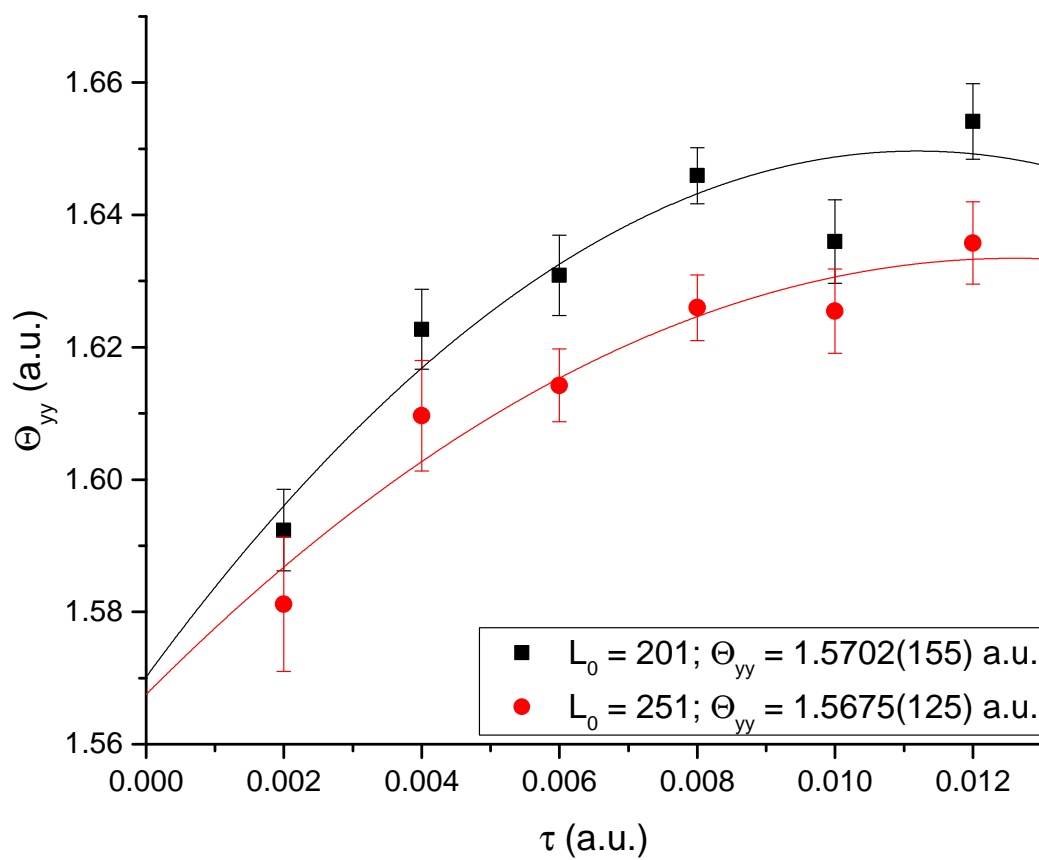
which is our Equation (25), where  $I$  is the number of samples (iterations) of  $E_{loc}$  obtained from a mixed distribution of electrons  $\mathbf{r}_i$ . This expression is advantageous since it does not require us to calculate the action of  $\hat{H}$  on  $\Phi_0$ , which we do not know.

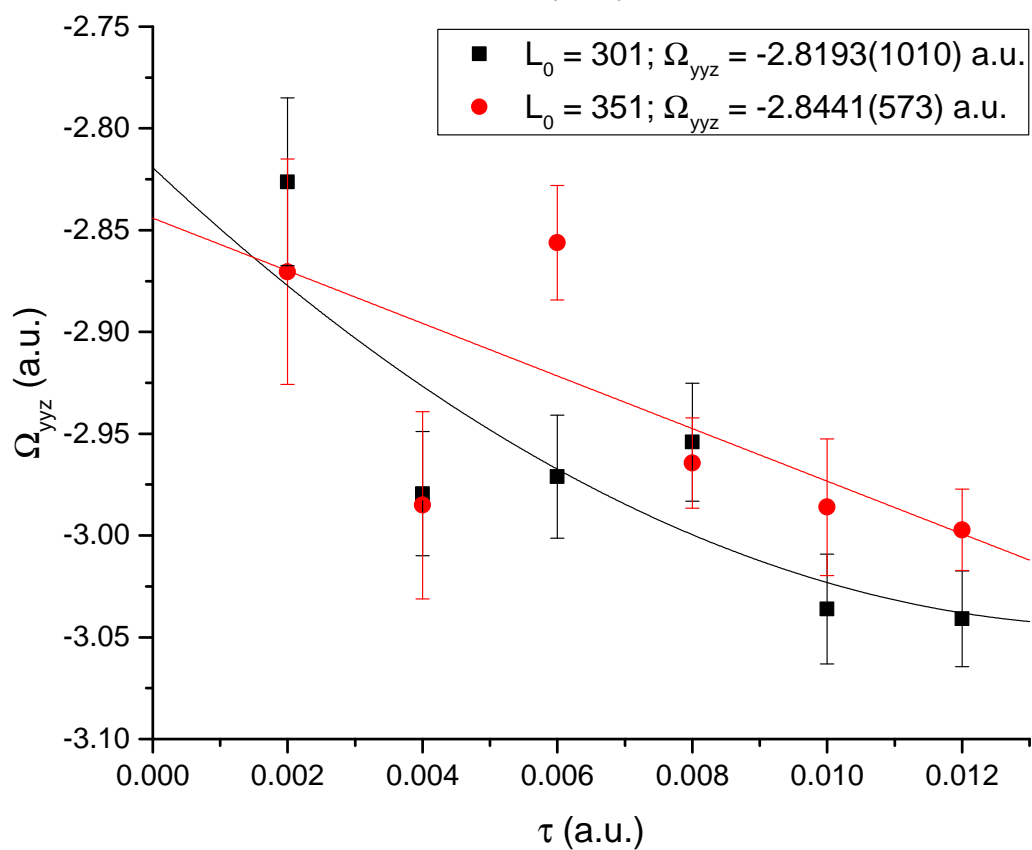
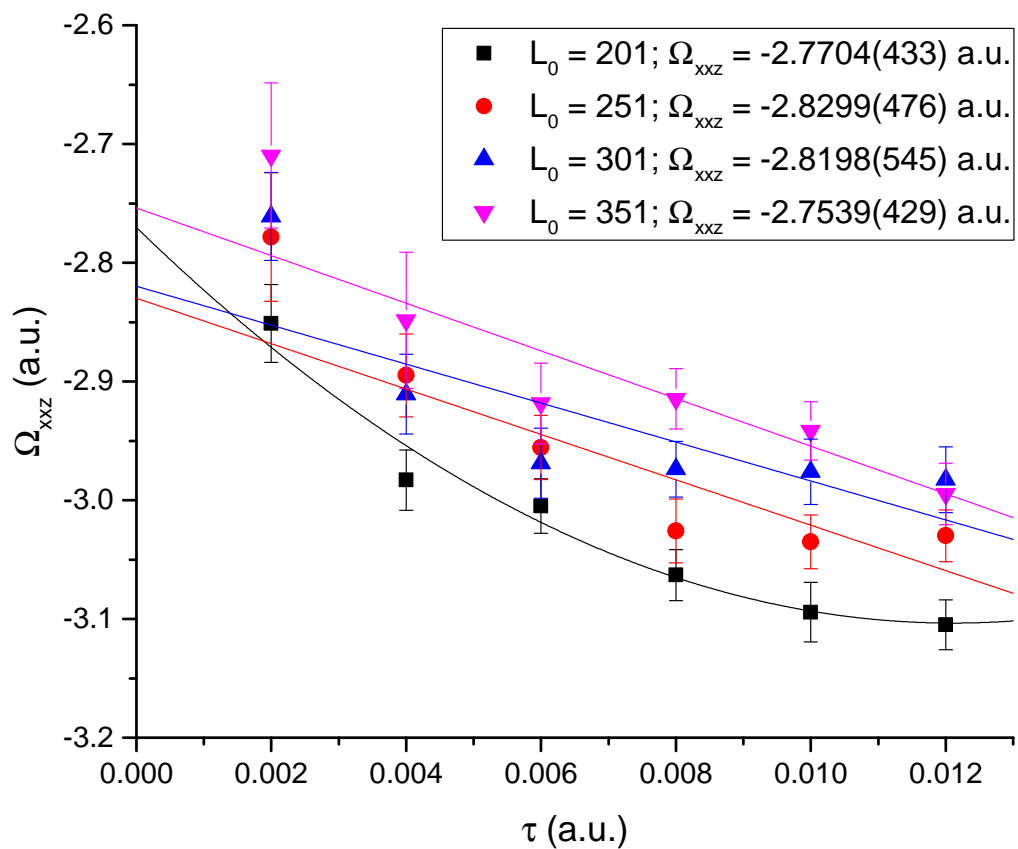
### A.3 Graphs of the Ground-state Energy and Electronic Properties of Lithium Hydride

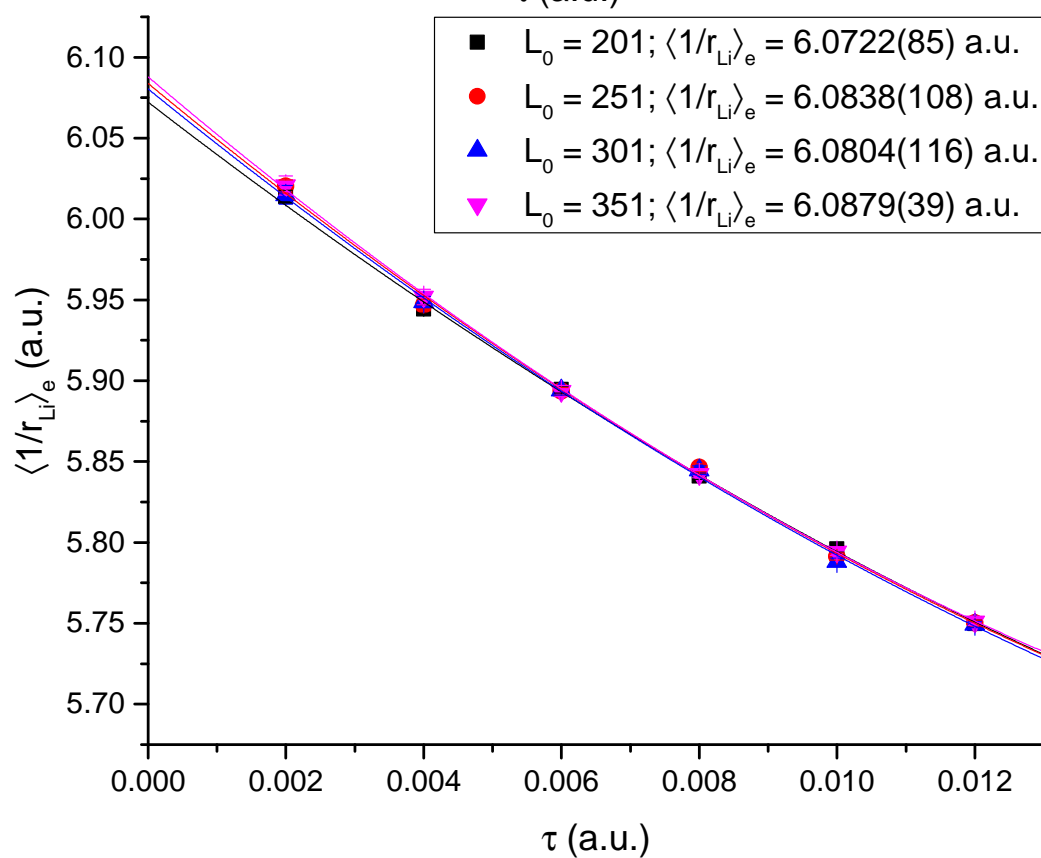
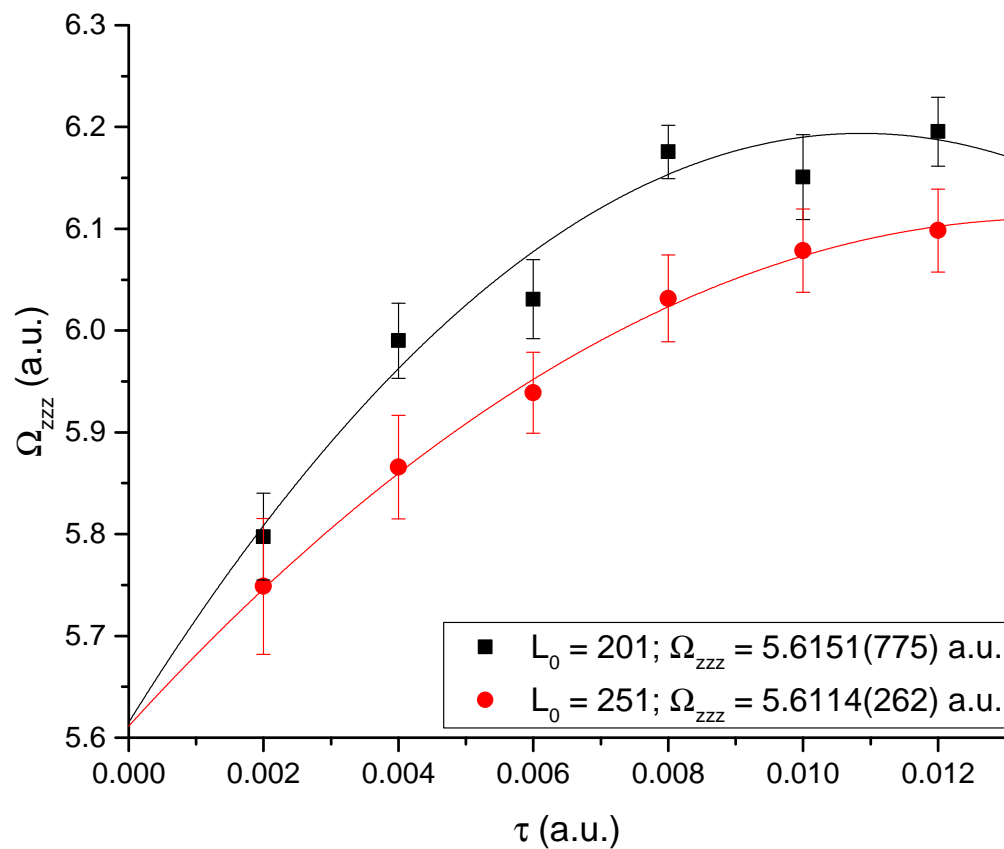


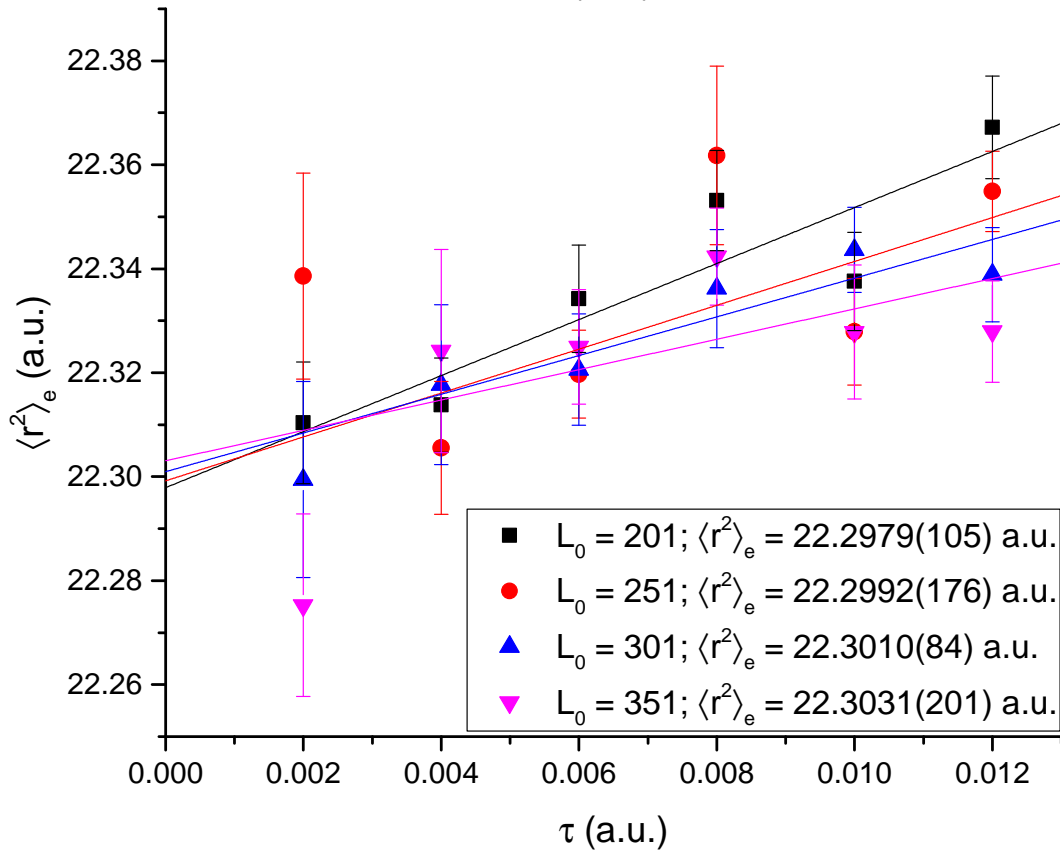
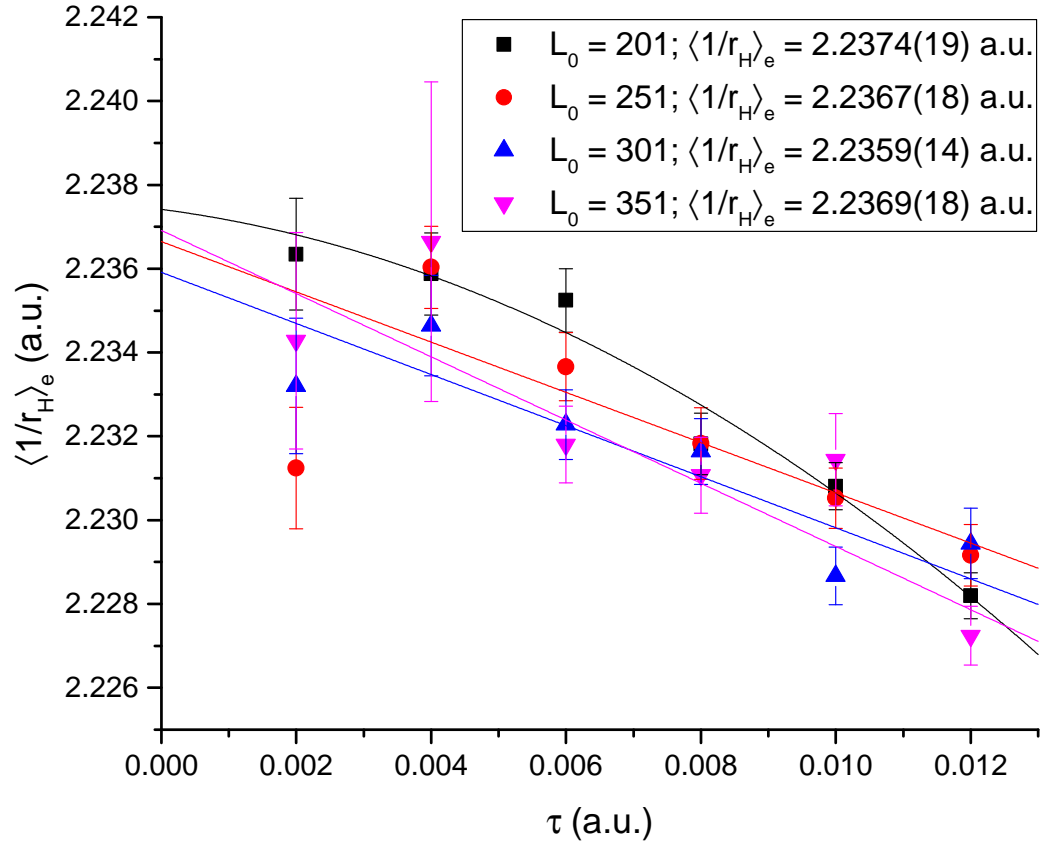
■	$L_0 = 41; E_{FN} = -8.07011(96) E_h$
●	$L_0 = 51; E_{FN} = -8.07060(57) E_h$
▲	$L_0 = 101; E_{FN} = -8.07043(119) E_h$
▼	$L_0 = 151; E_{FN} = -8.07093(90) E_h$
◆	$L_0 = 201; E_{FN} = -8.07085(38) E_h$
◄	$L_0 = 251; E_{FN} = -8.07033(140) E_h$
►	$L_0 = 301; E_{FN} = -8.07074(152) E_h$



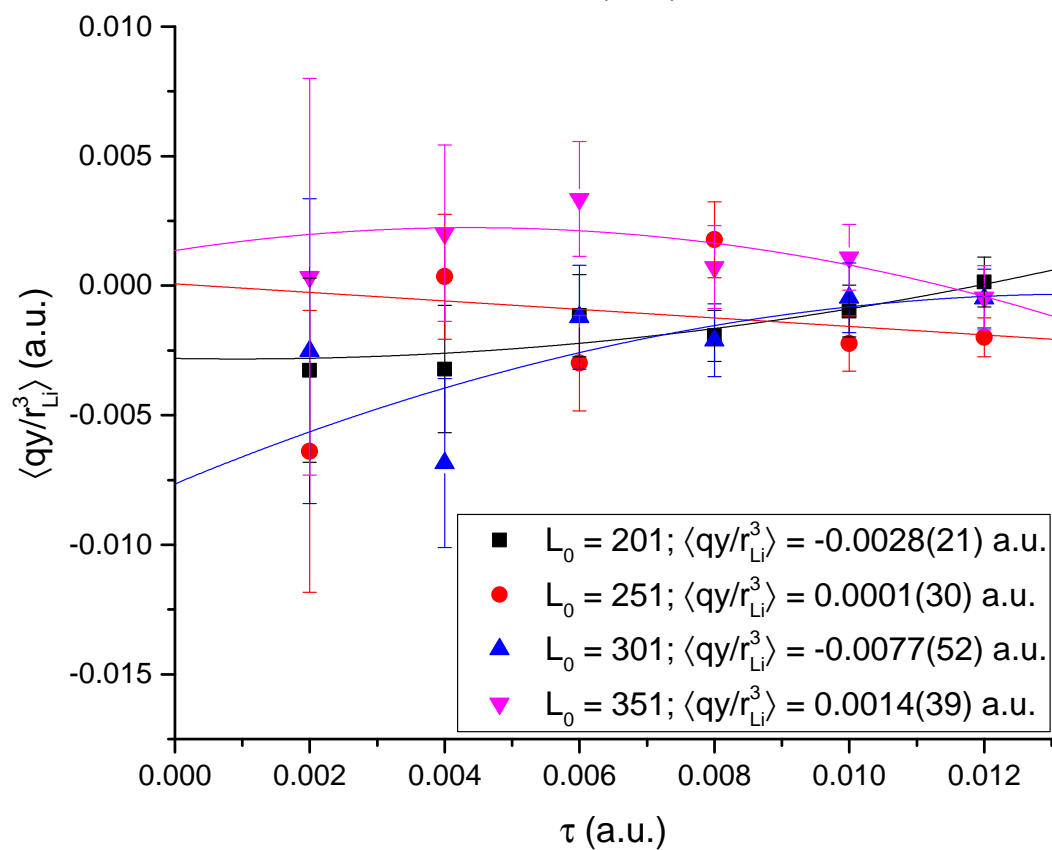
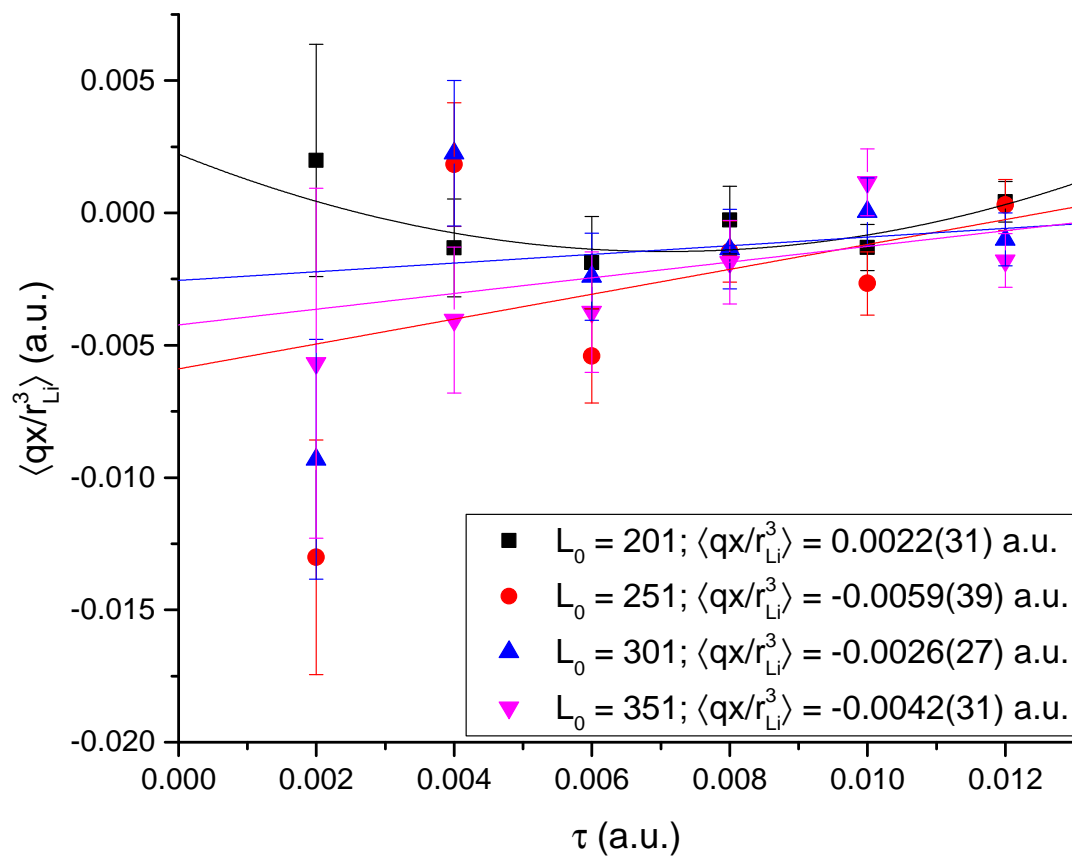


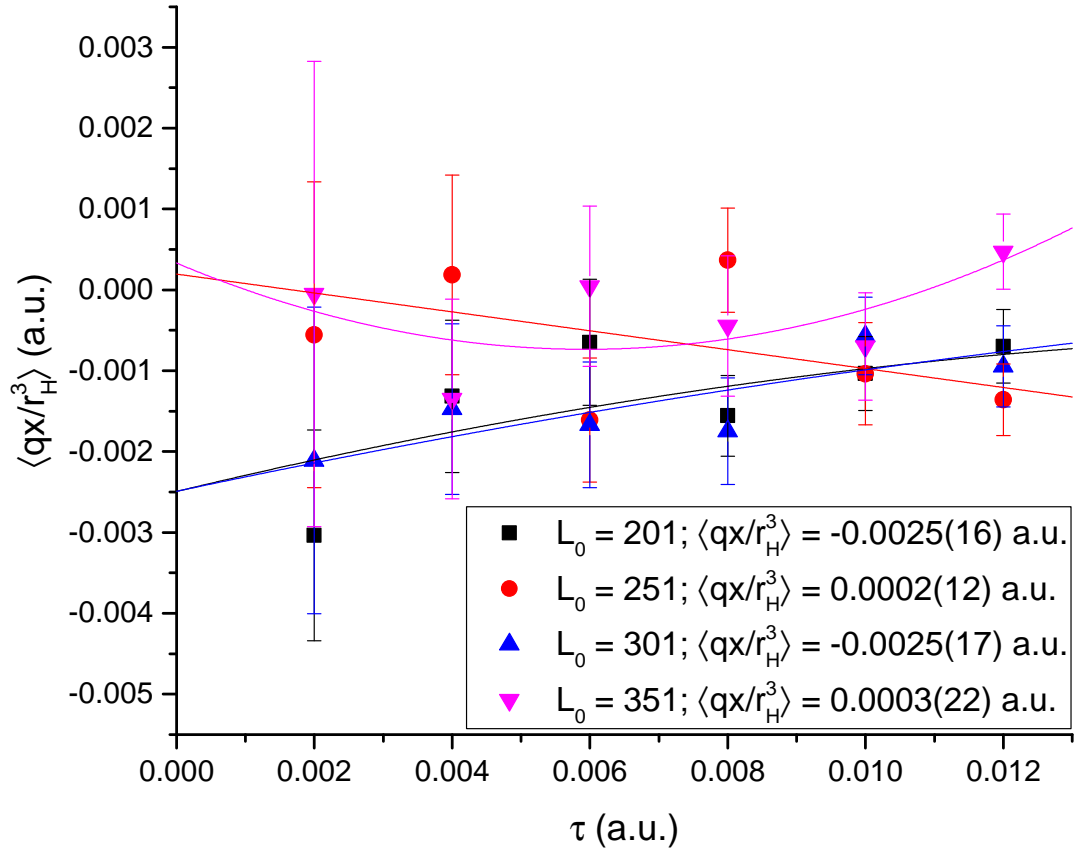
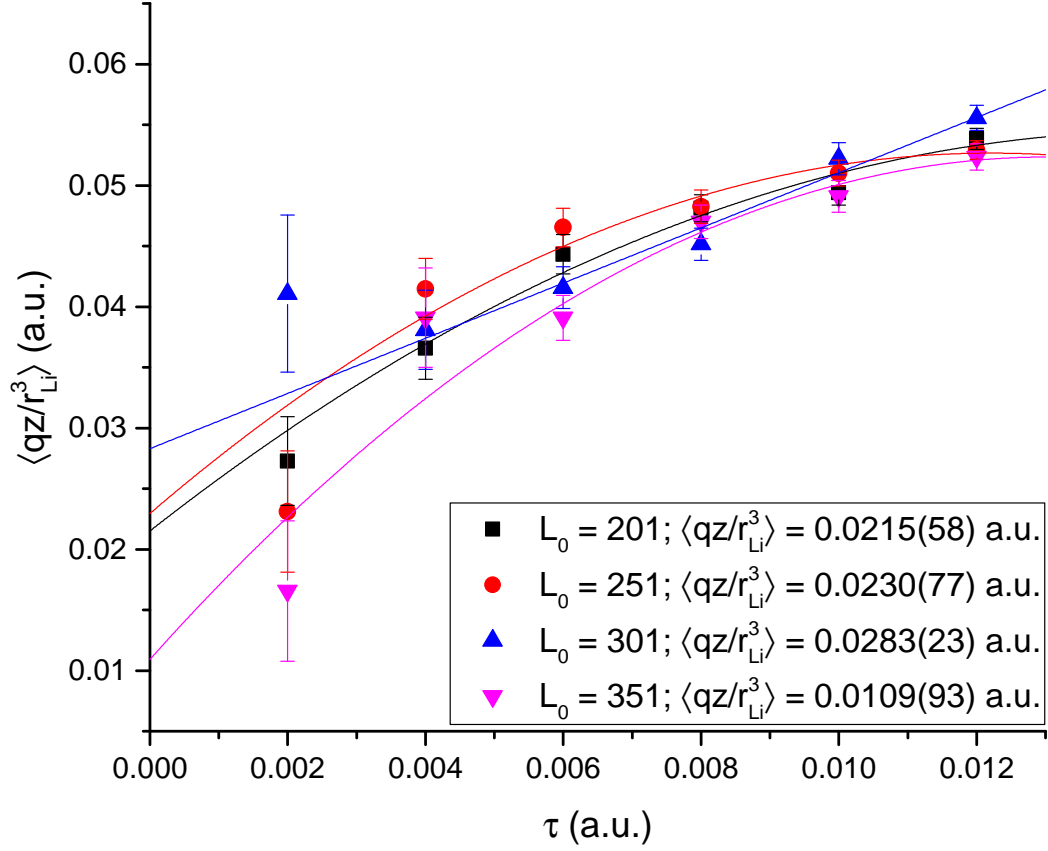


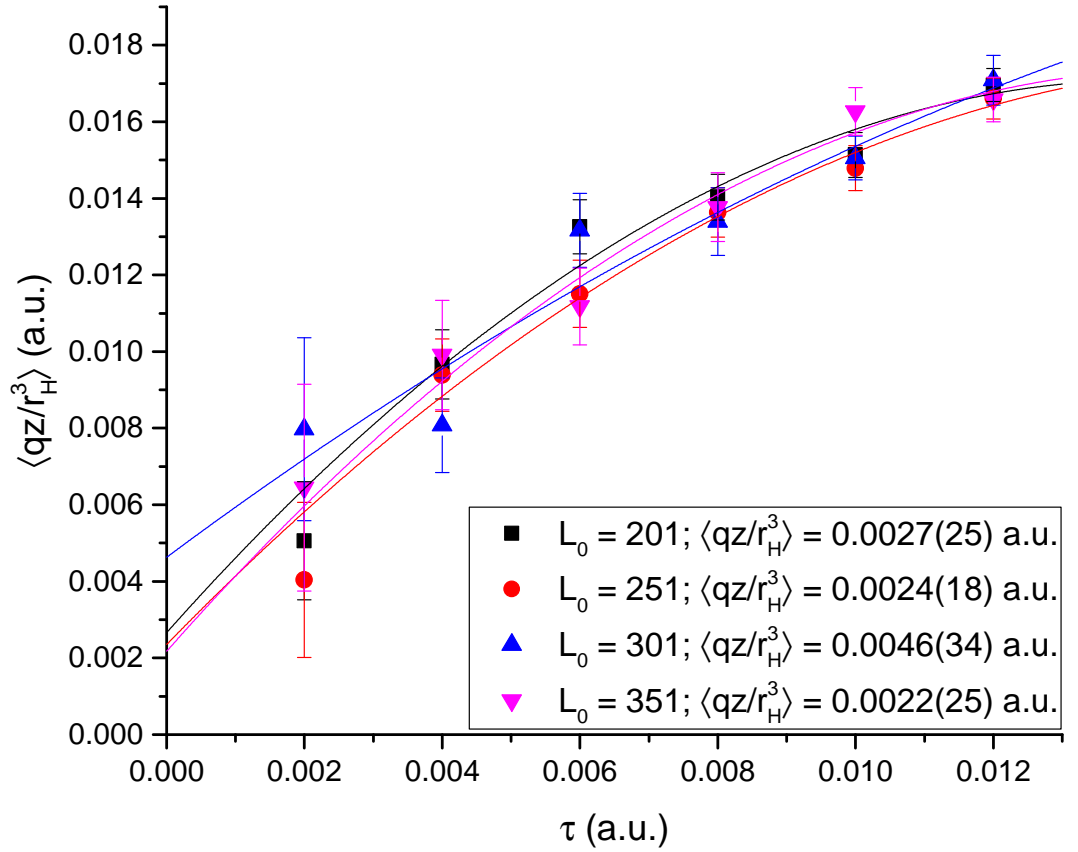
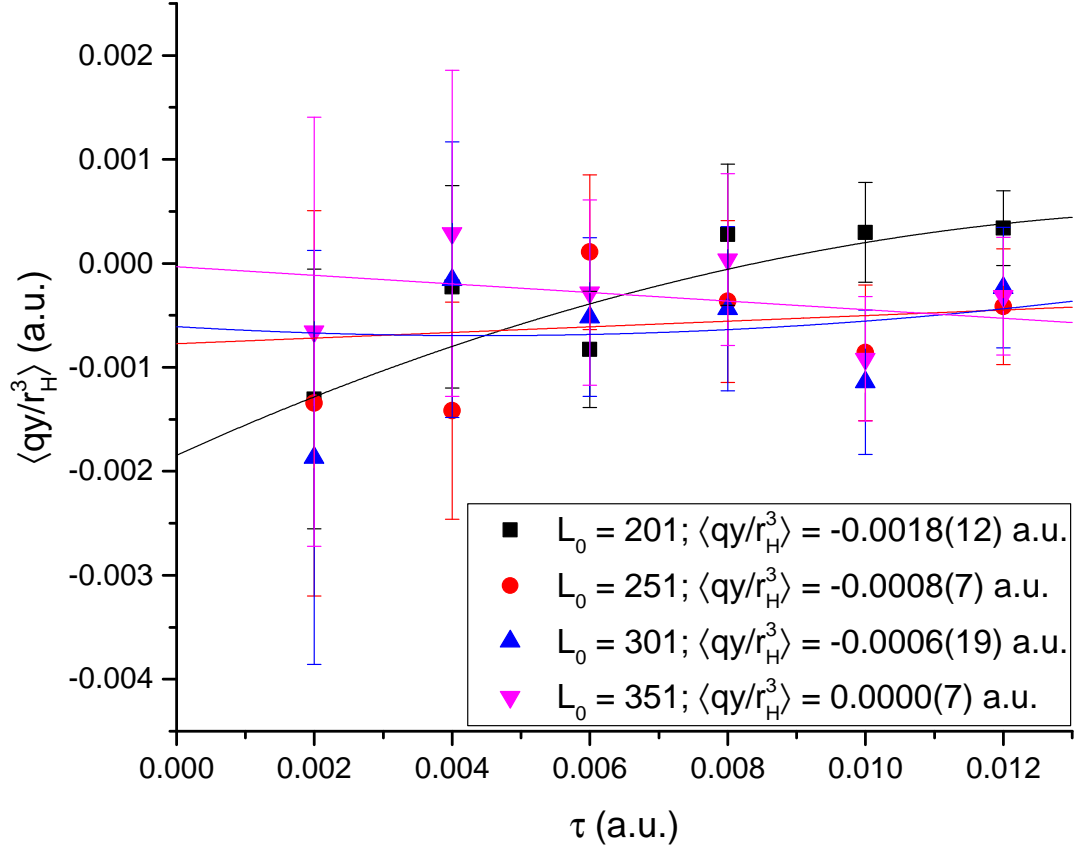


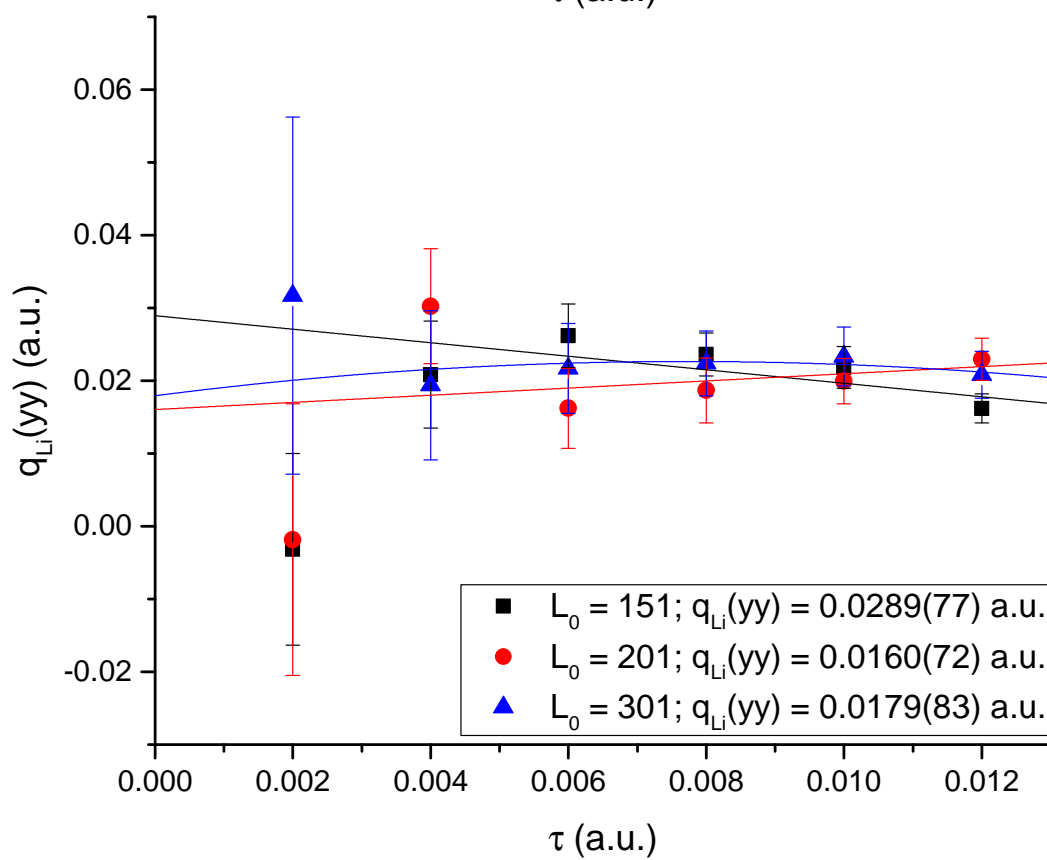
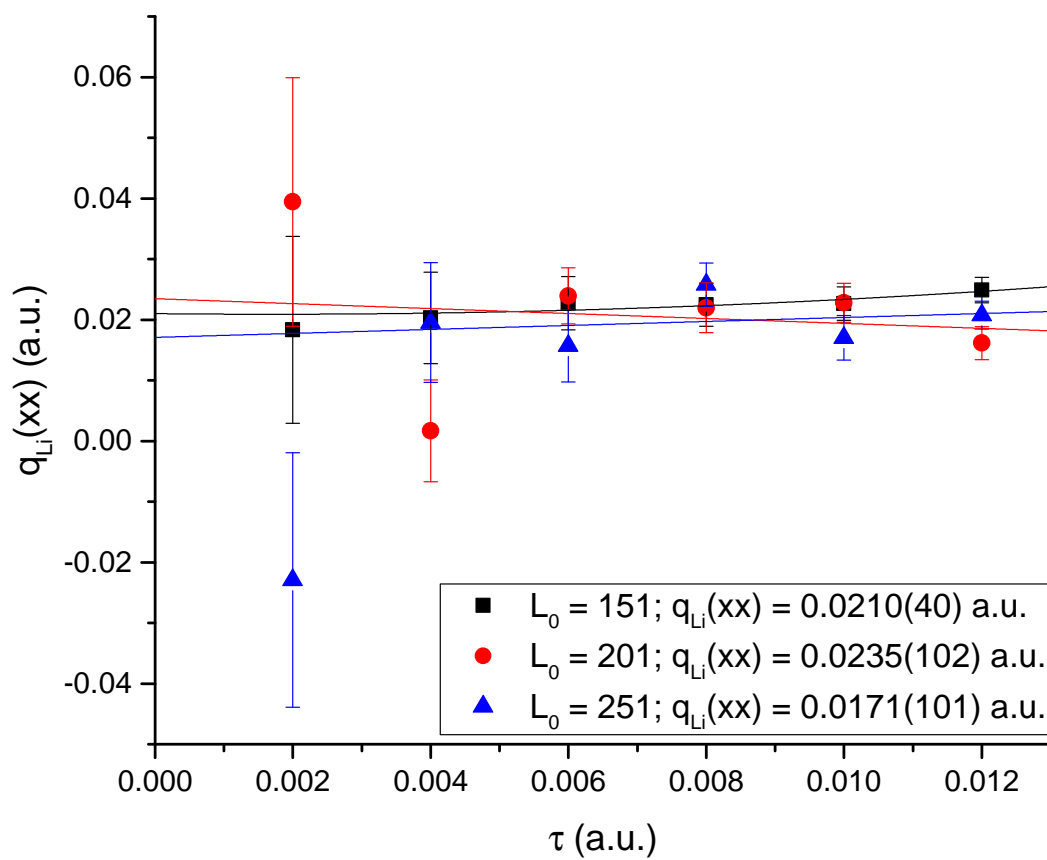


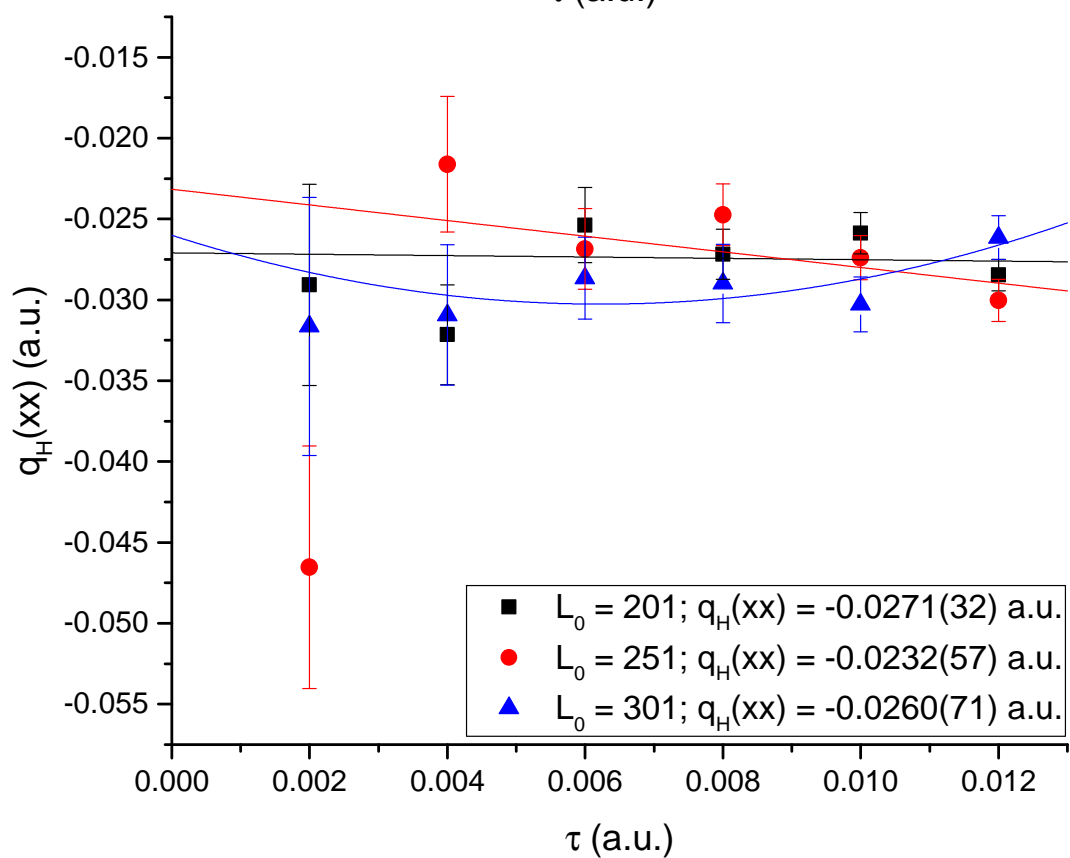
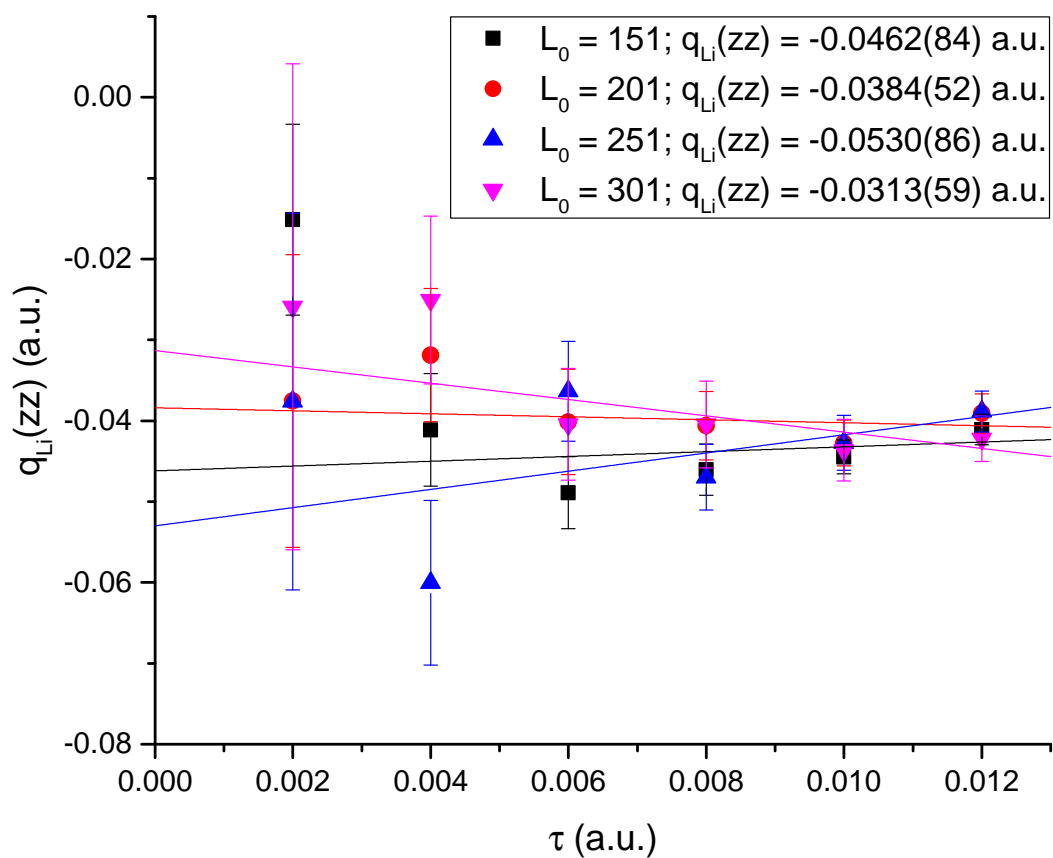


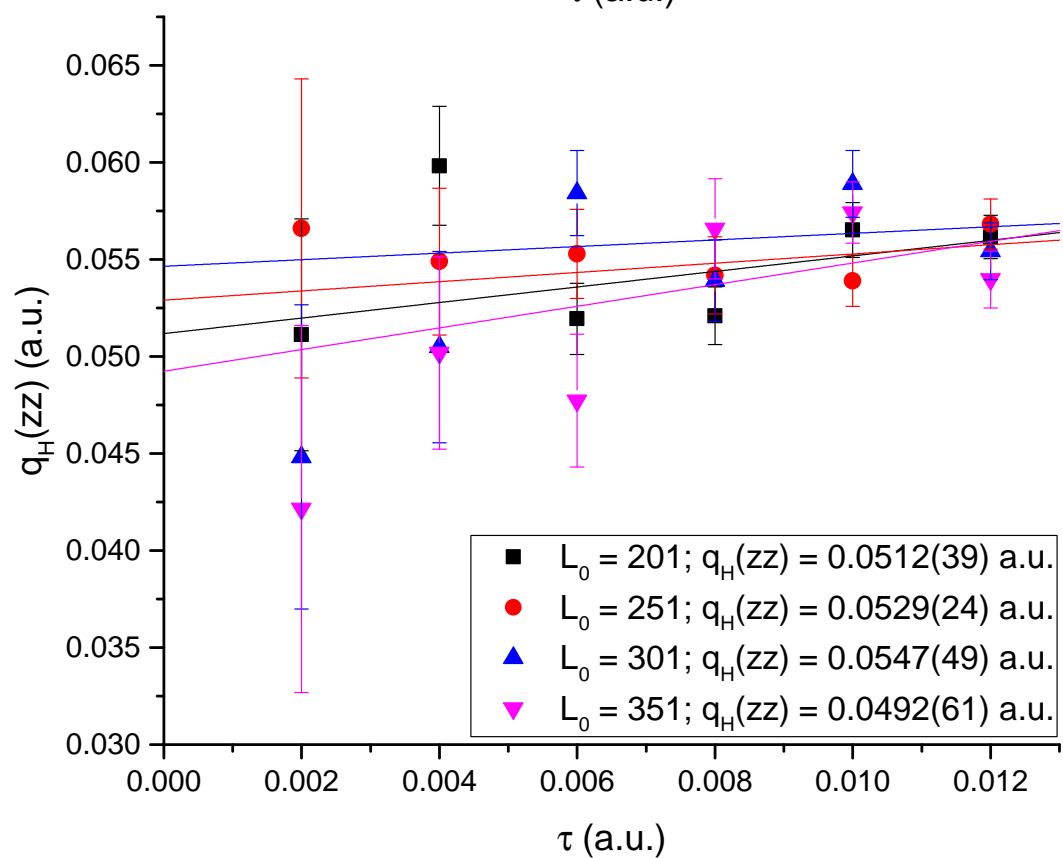
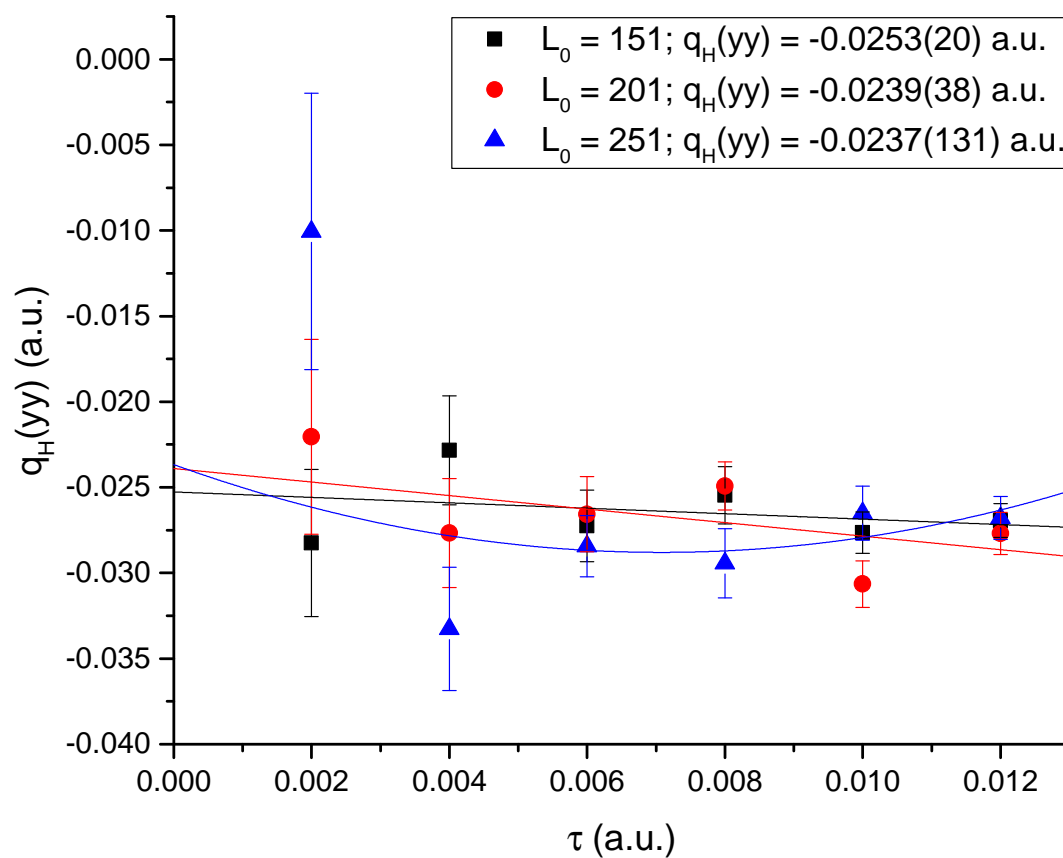


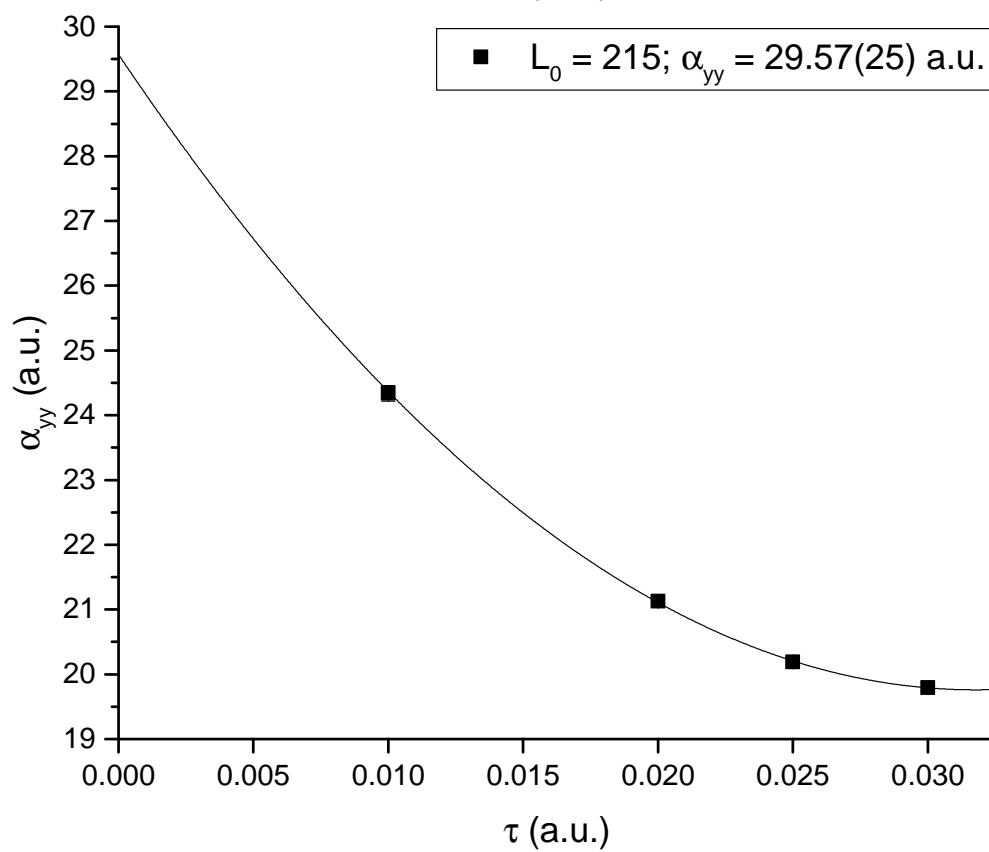
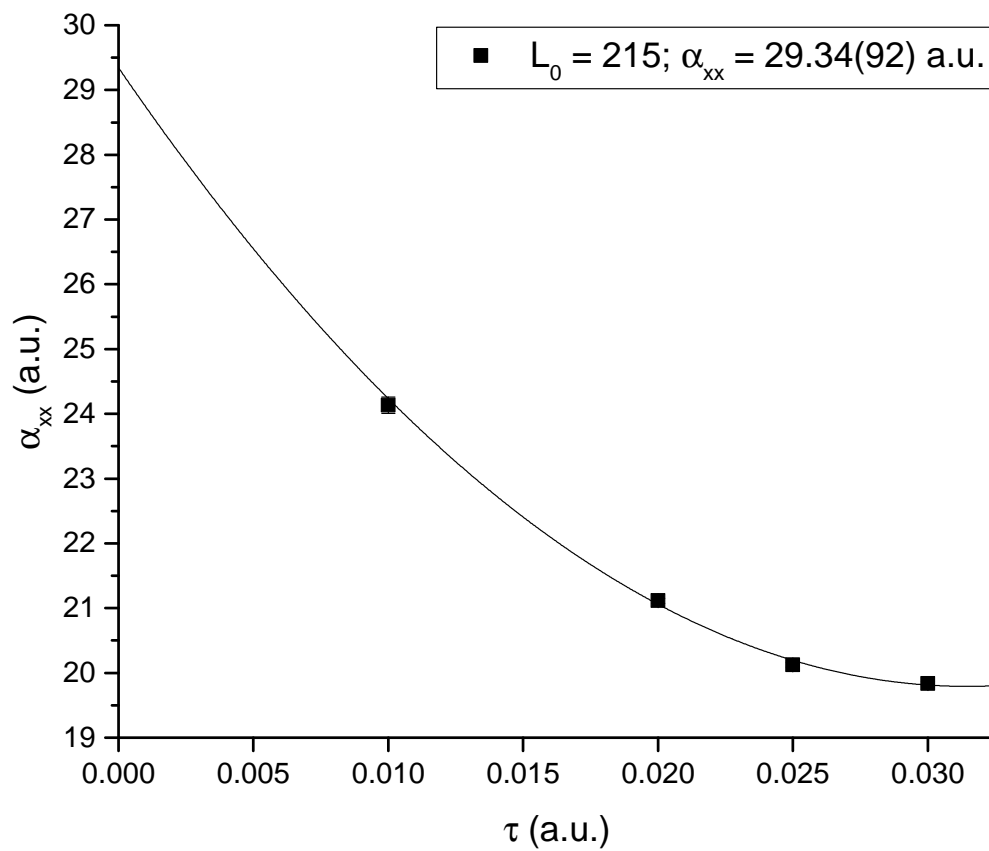


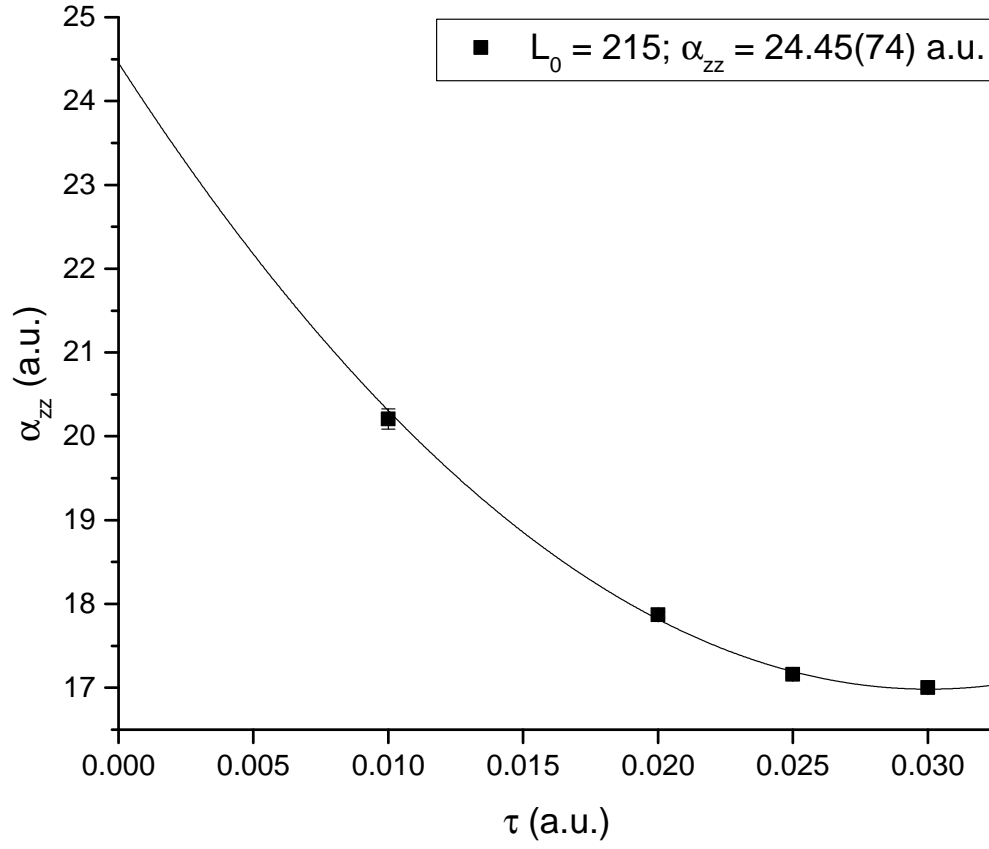






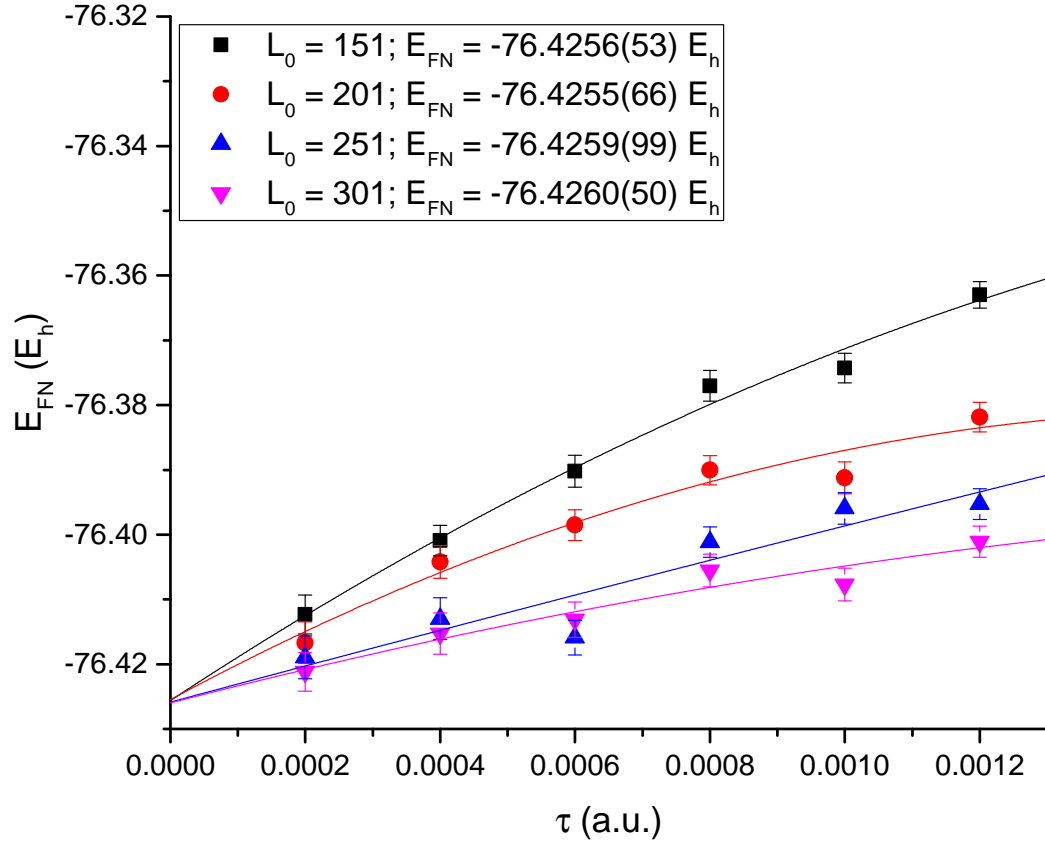


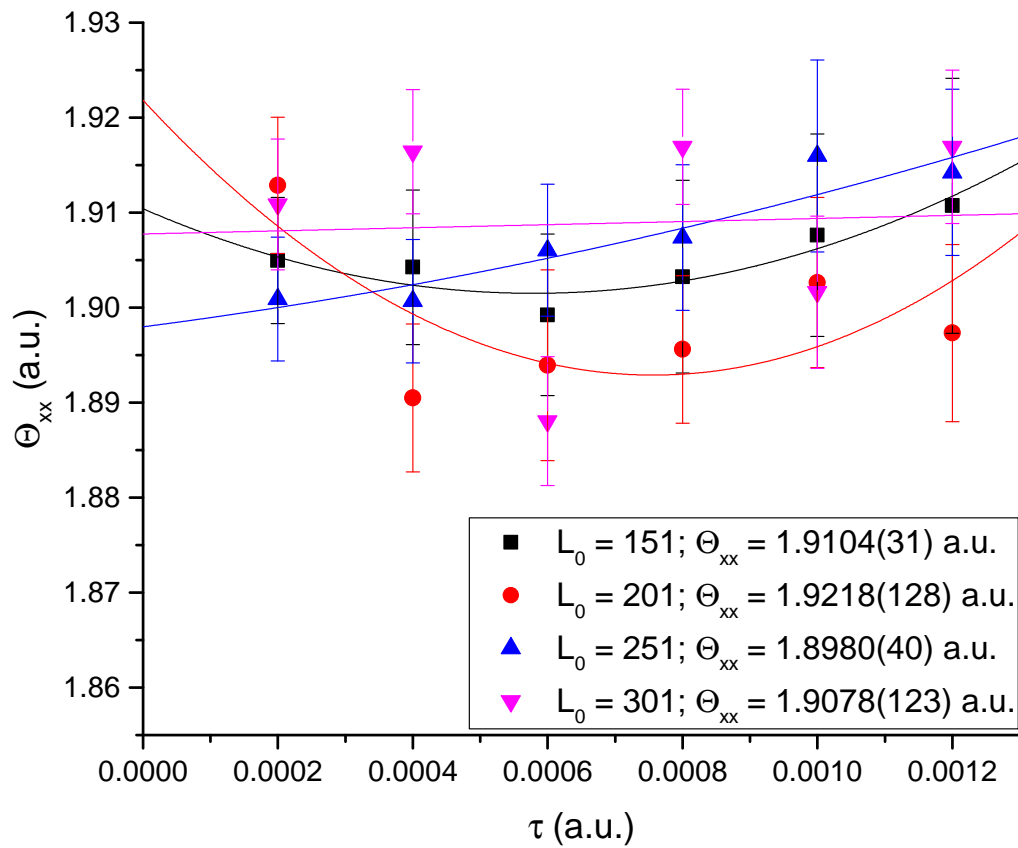
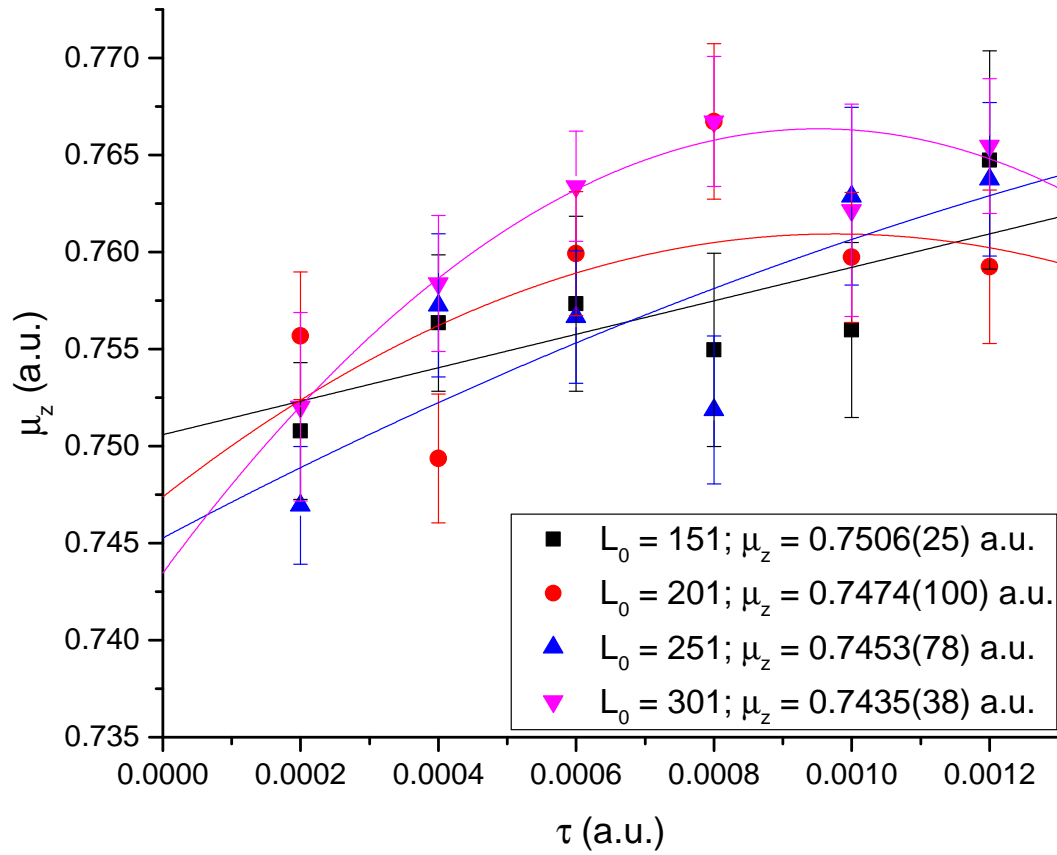


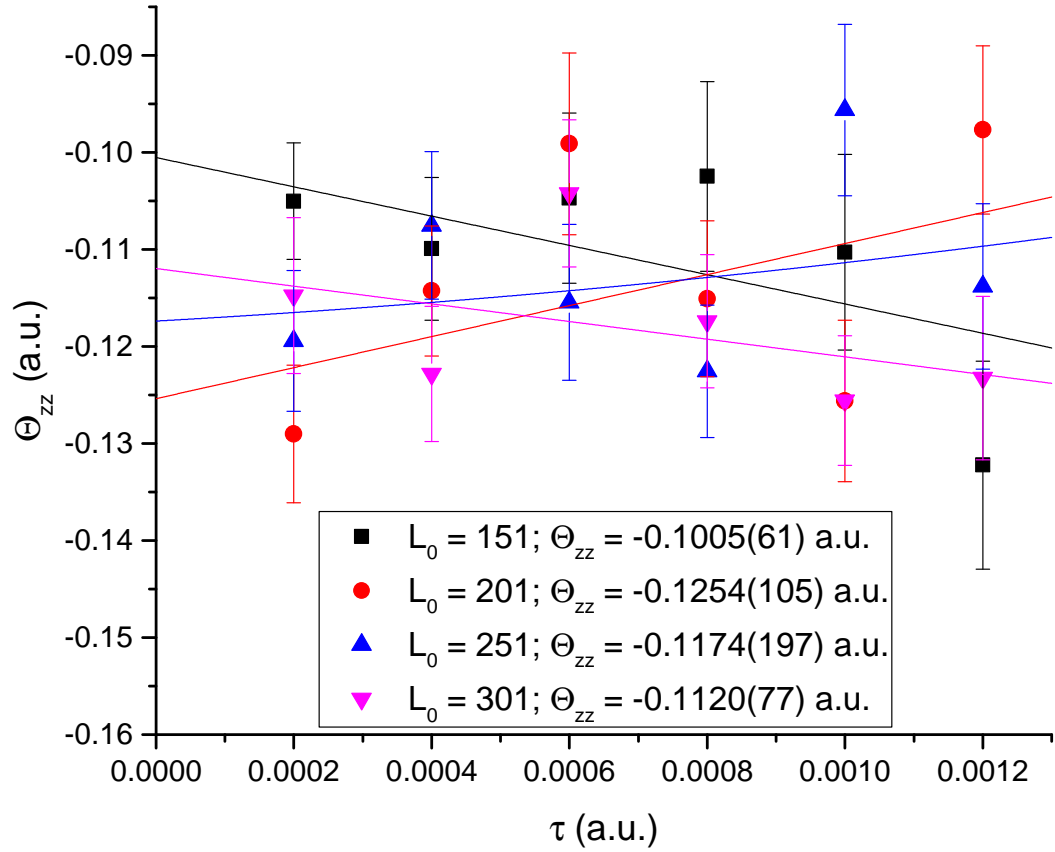
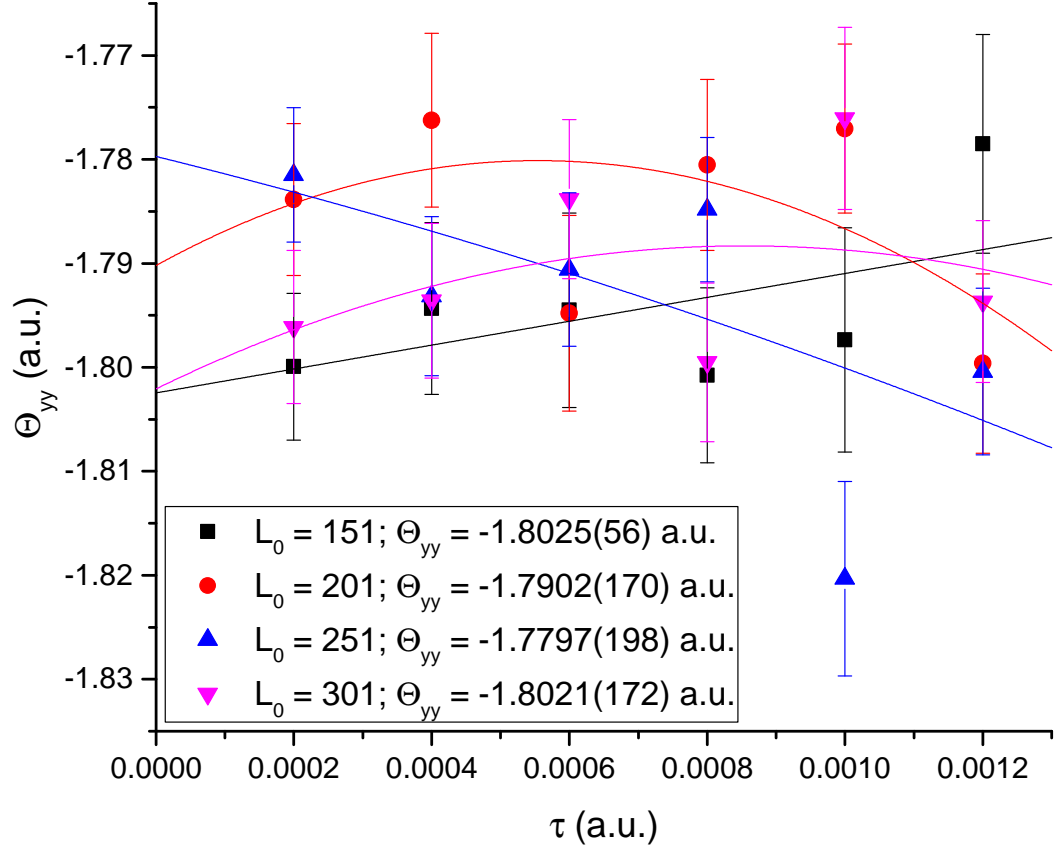


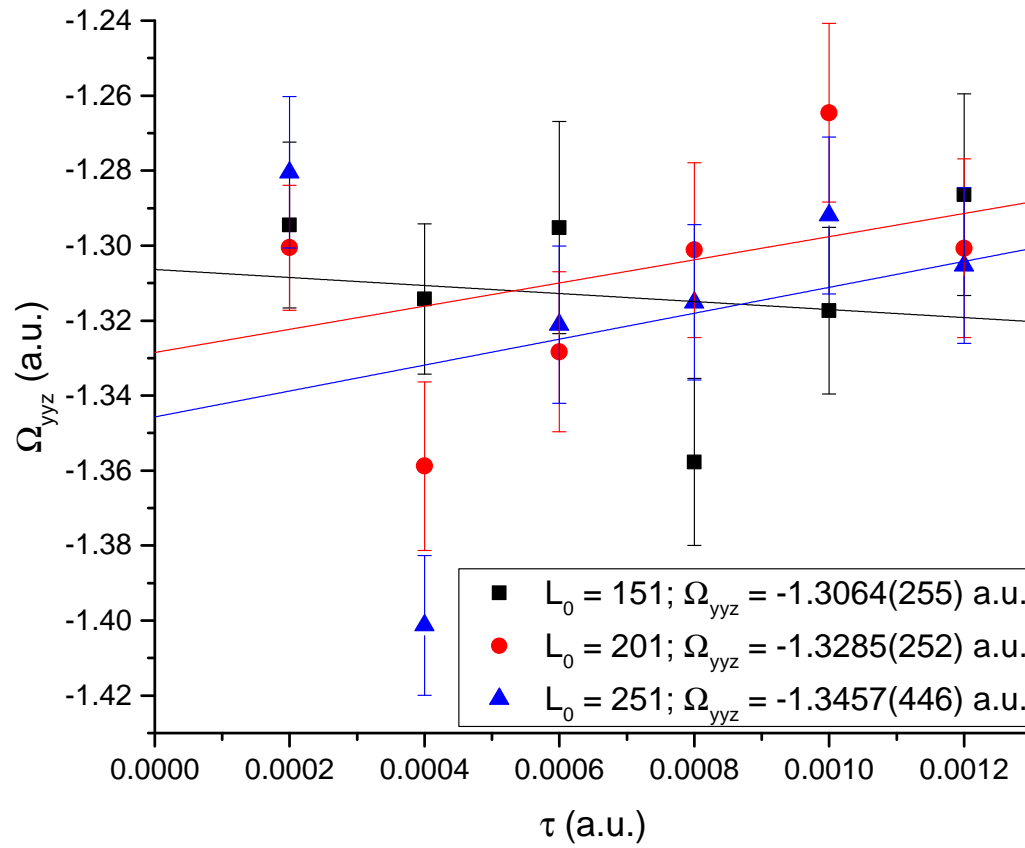
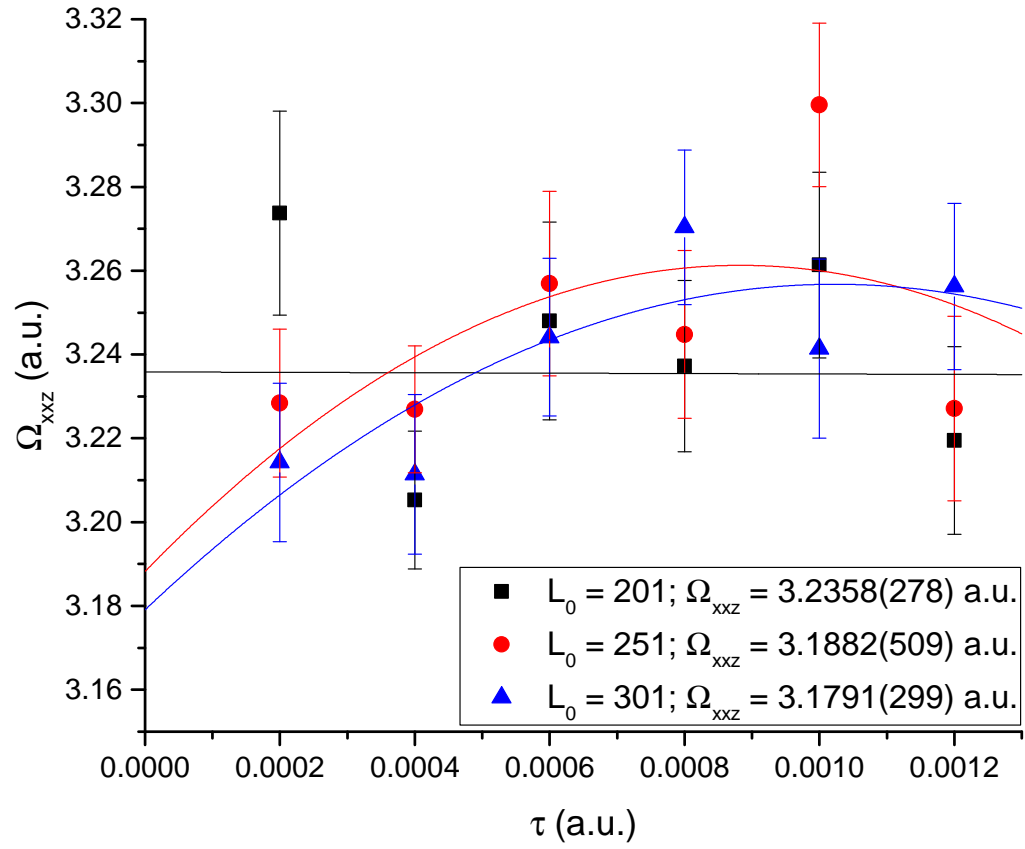


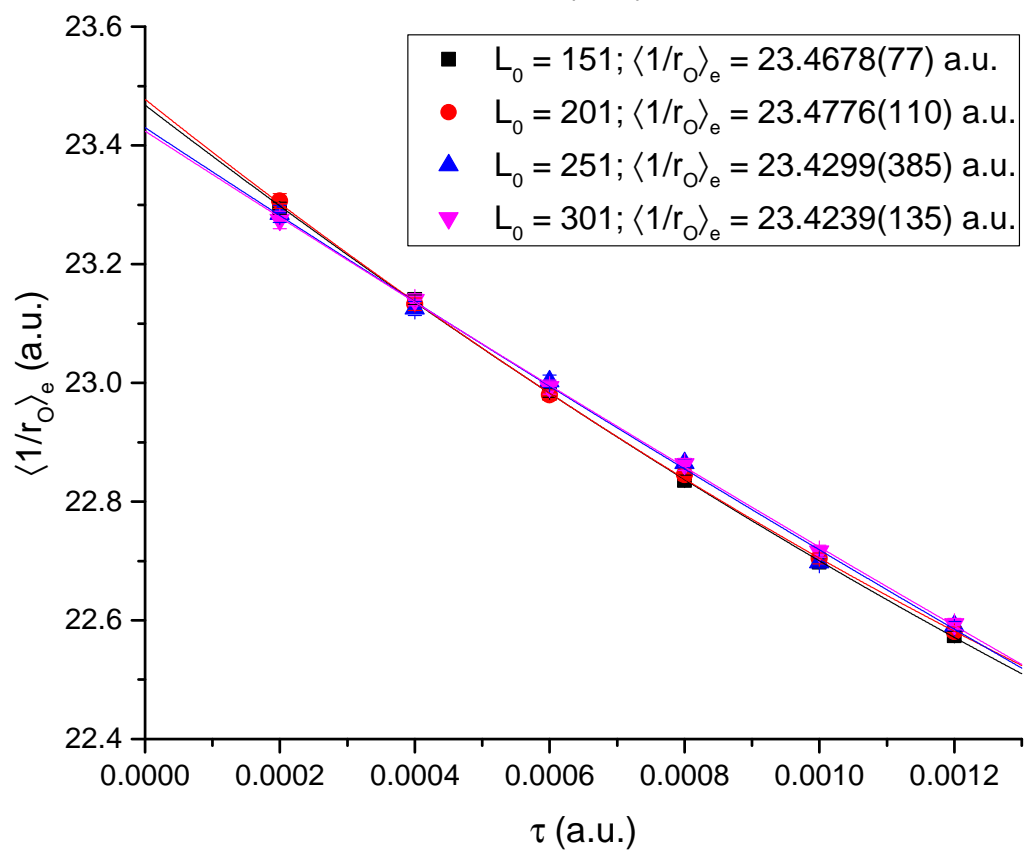
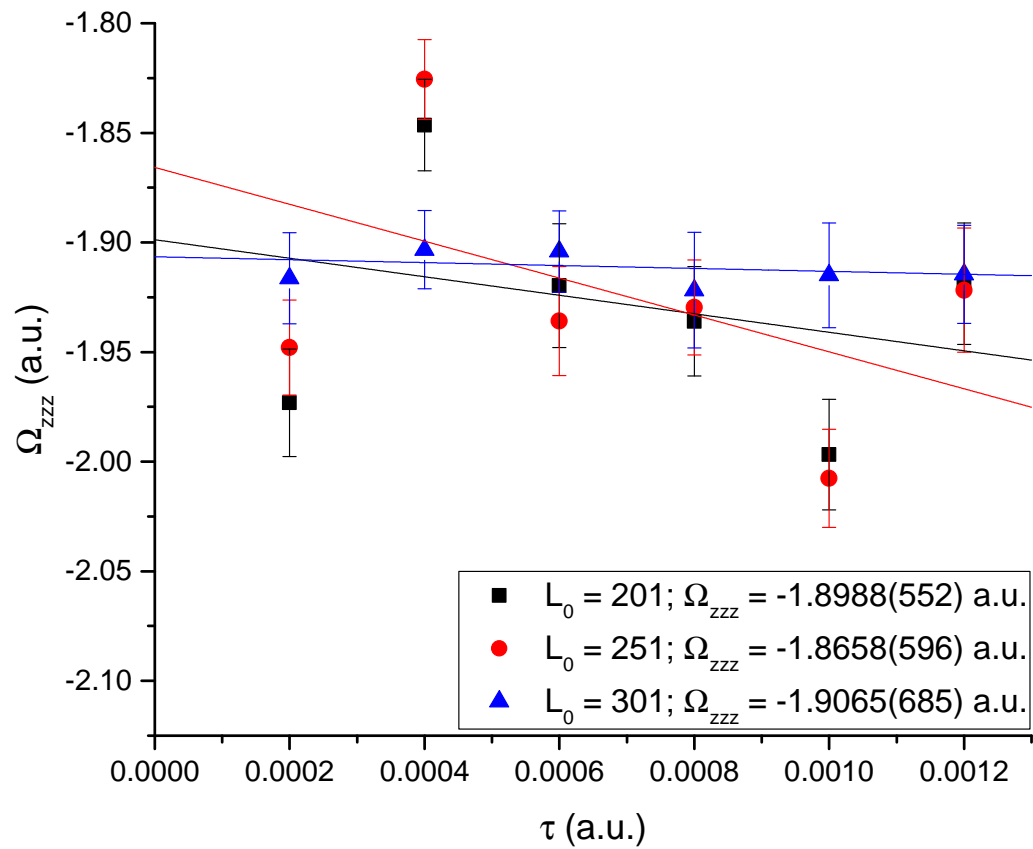
### A.4 Graphs of the Ground-state Energy and Electronic Properties of Water

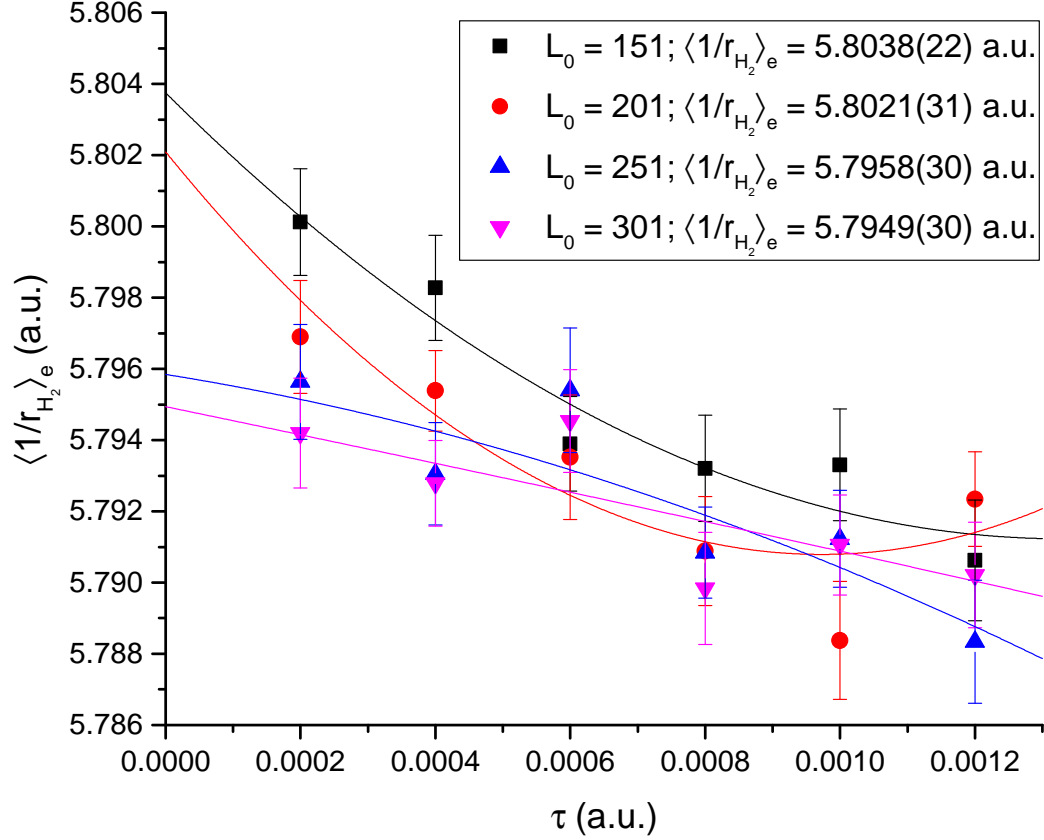
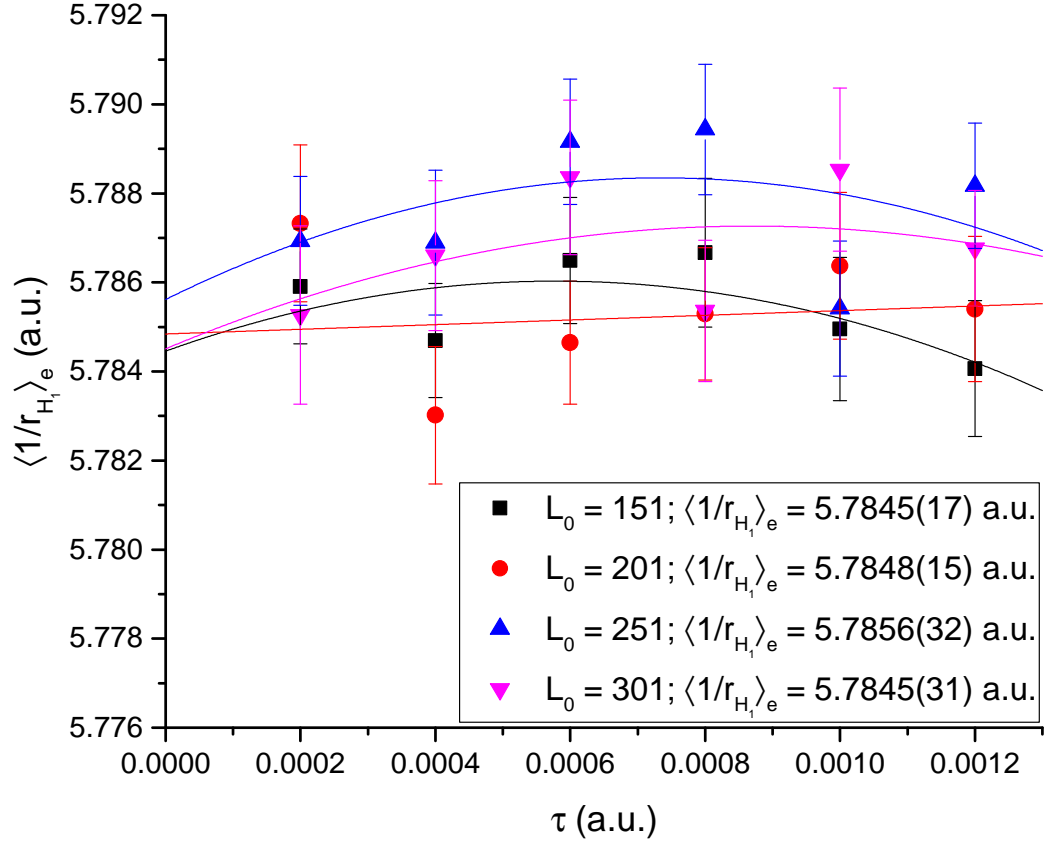


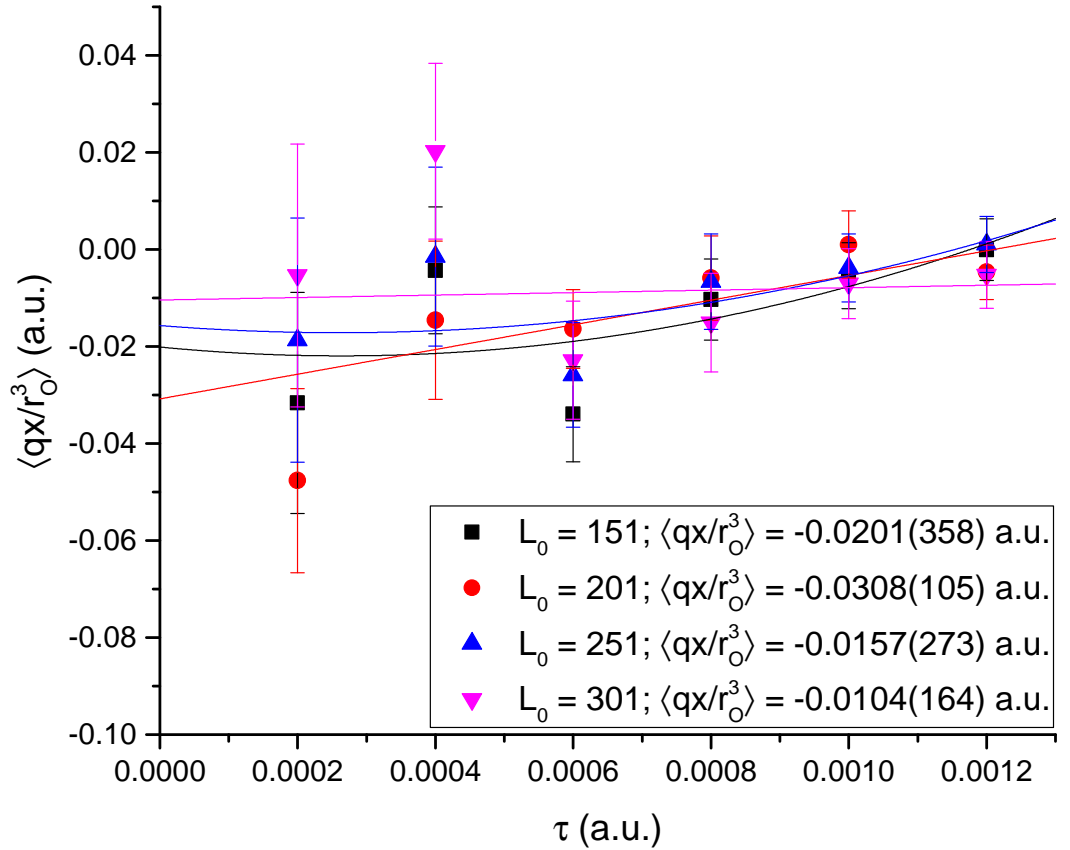
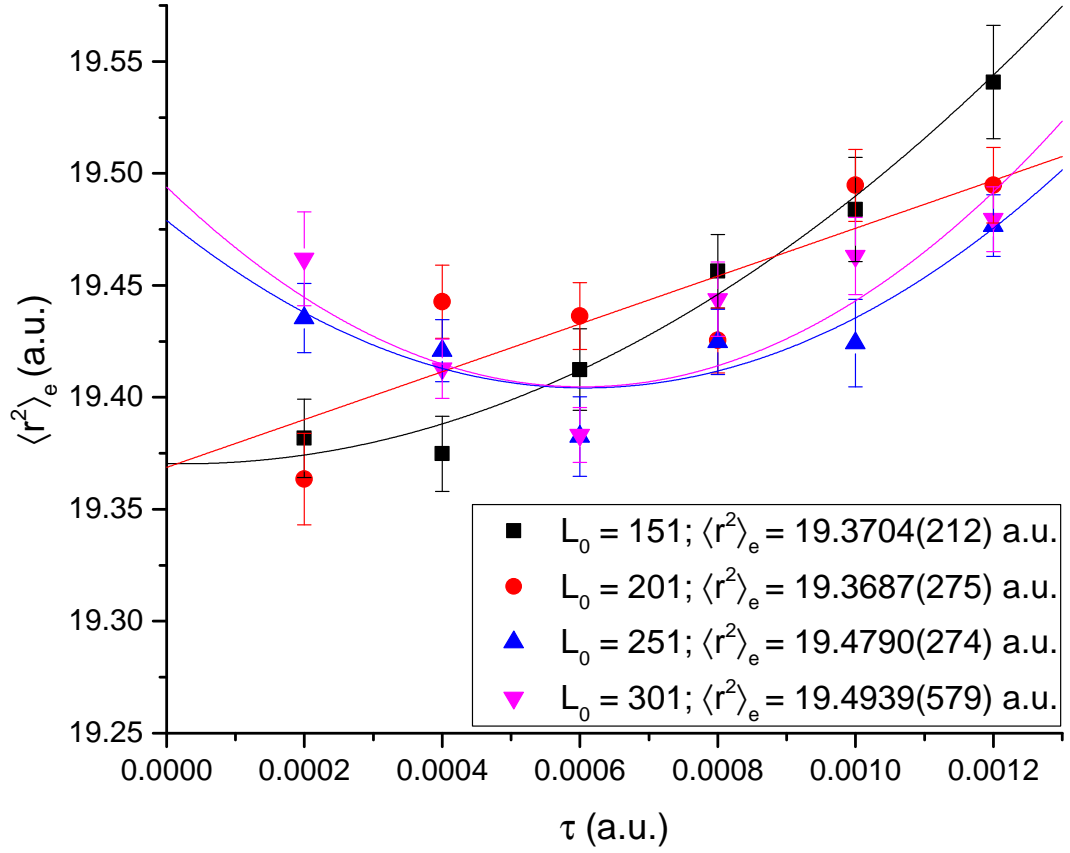


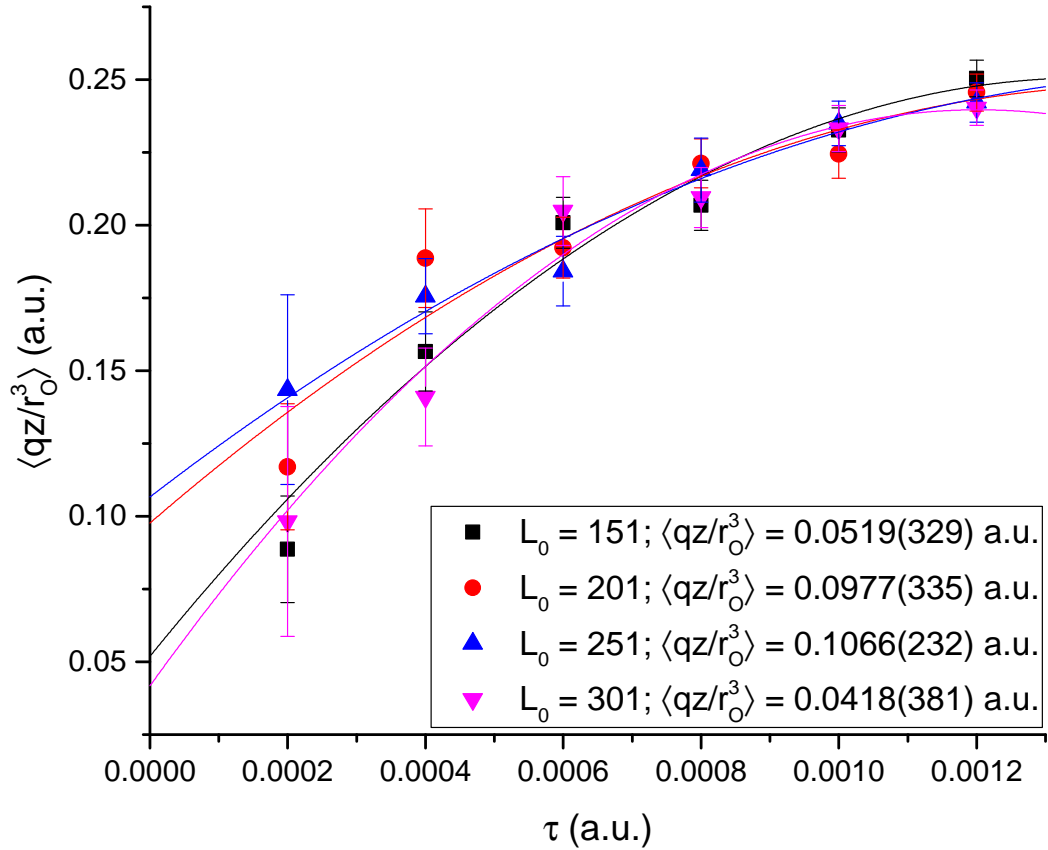
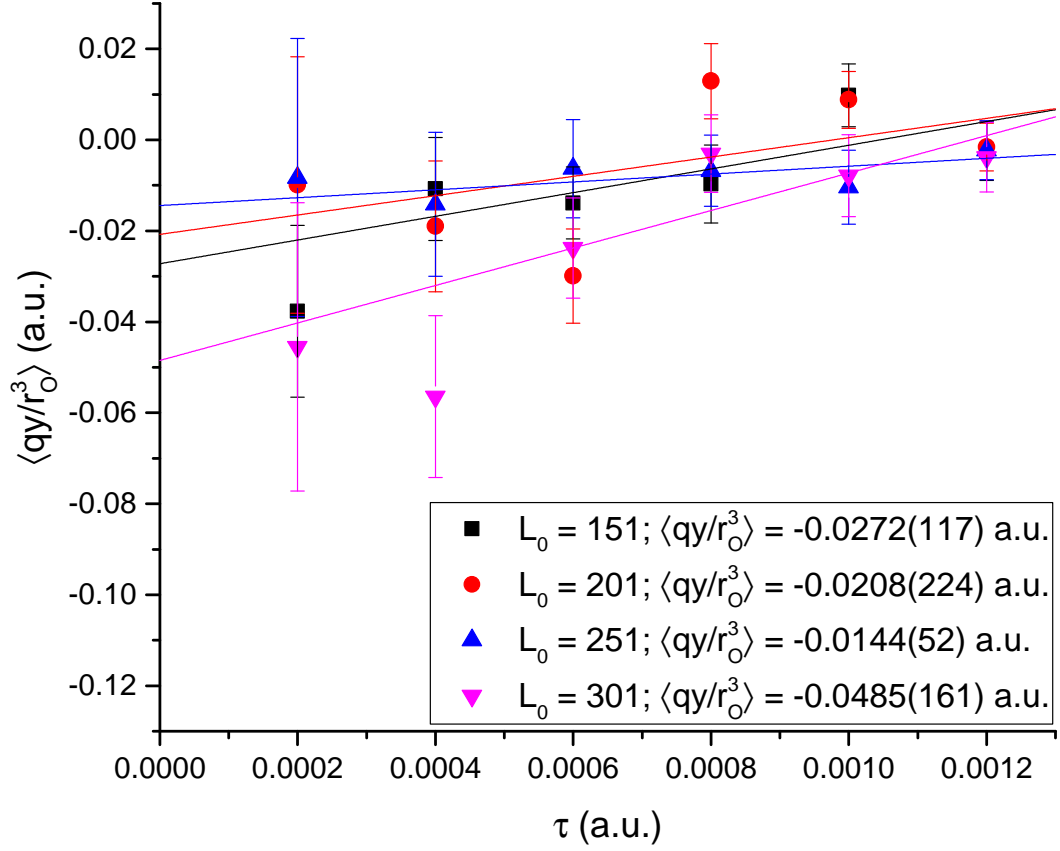




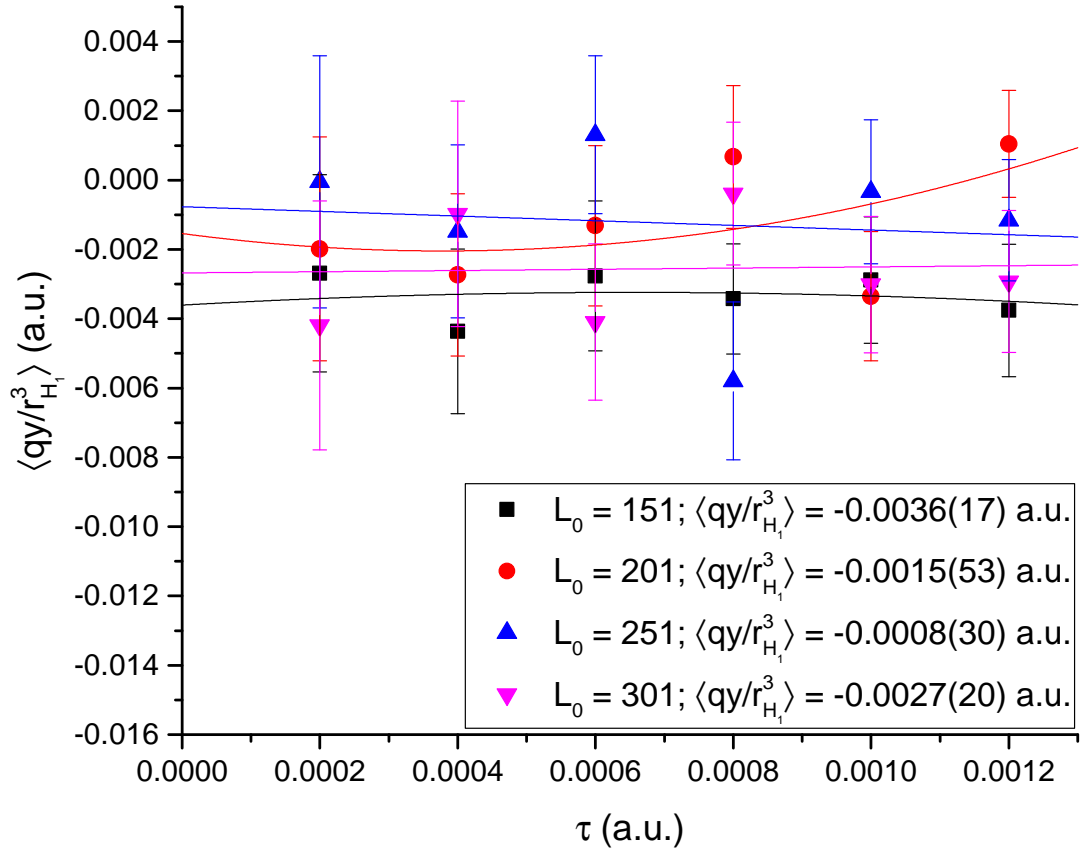
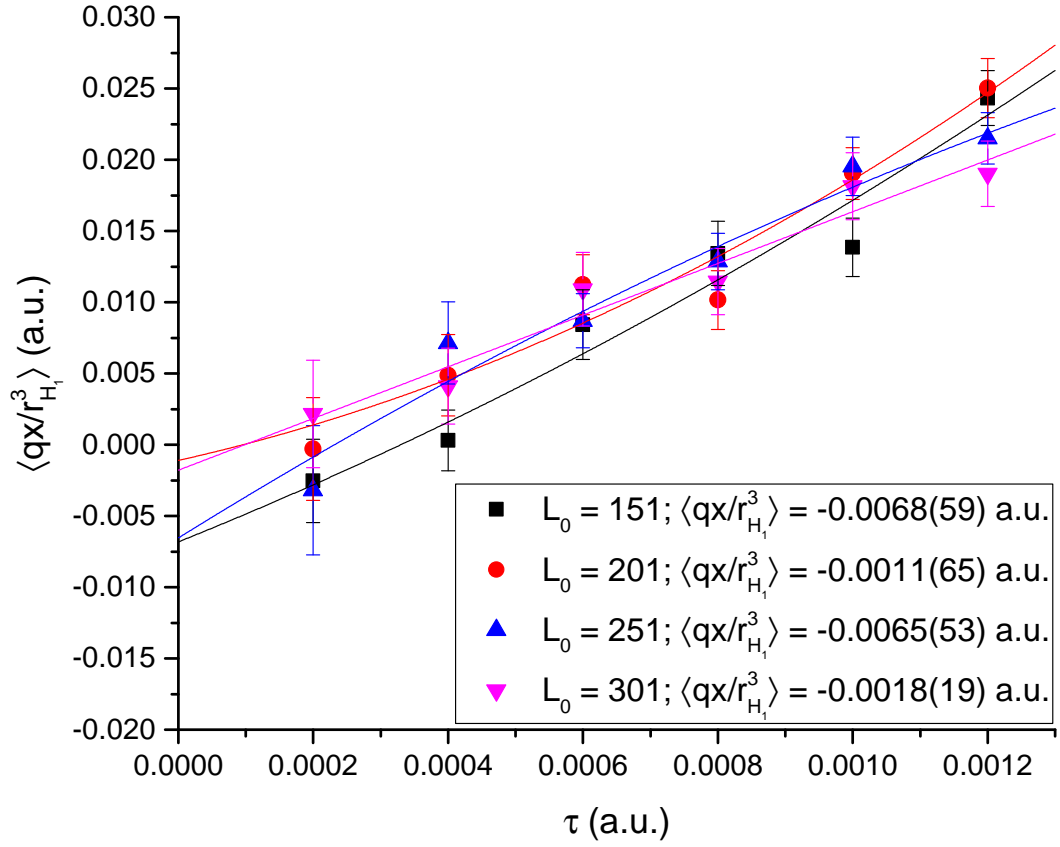


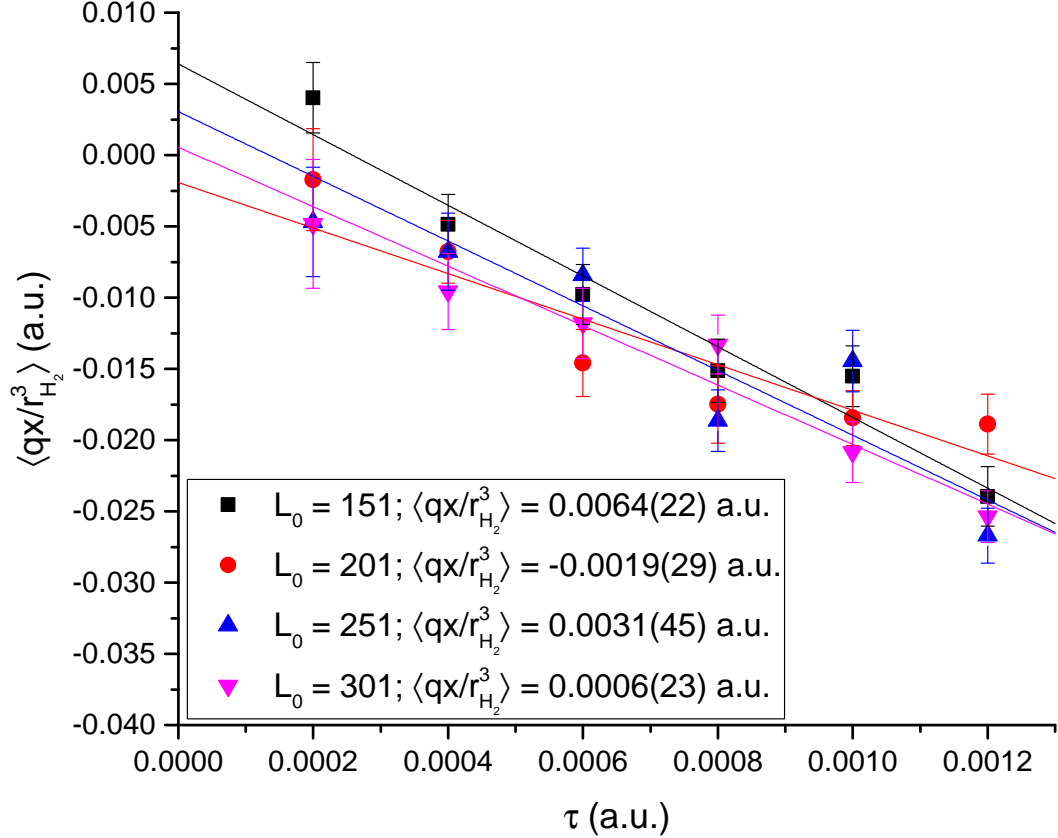
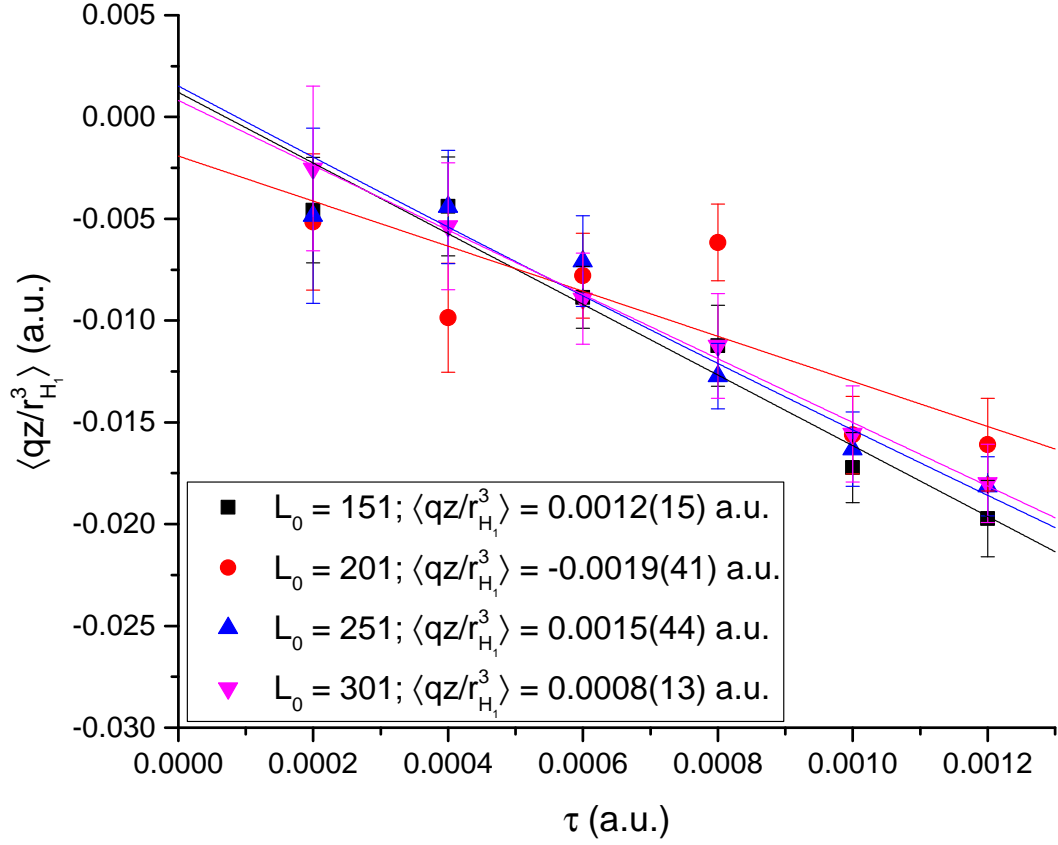


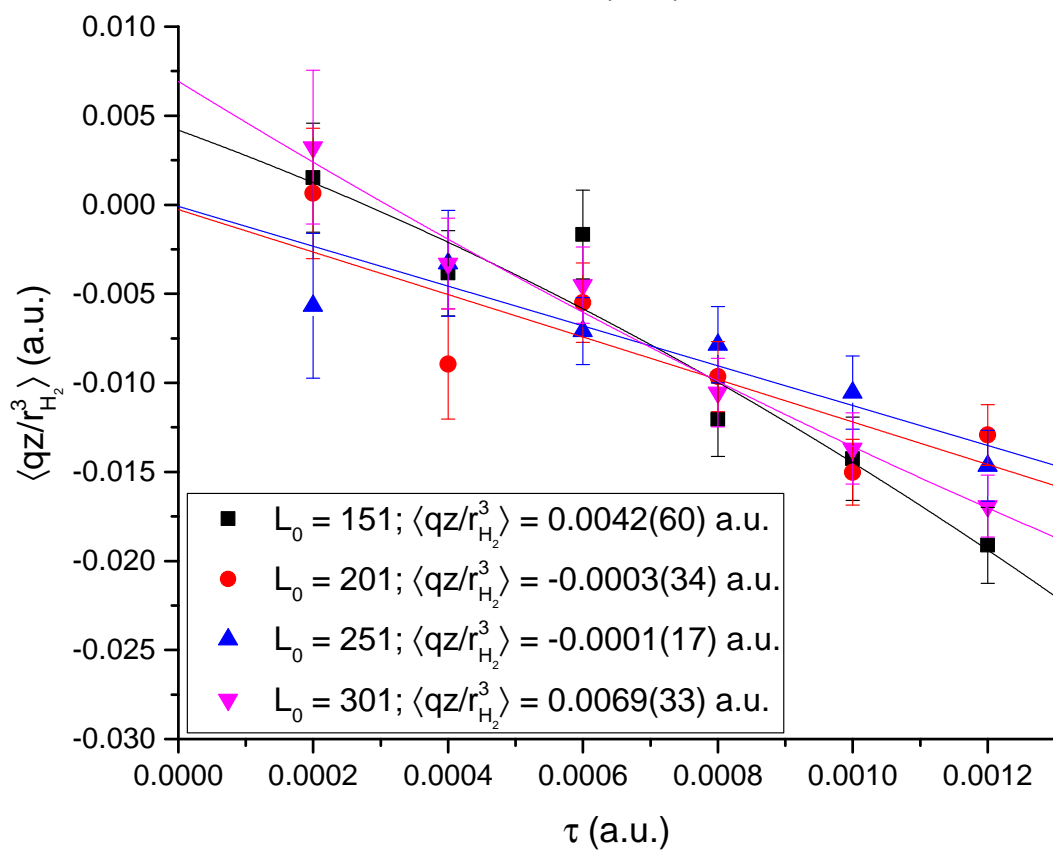
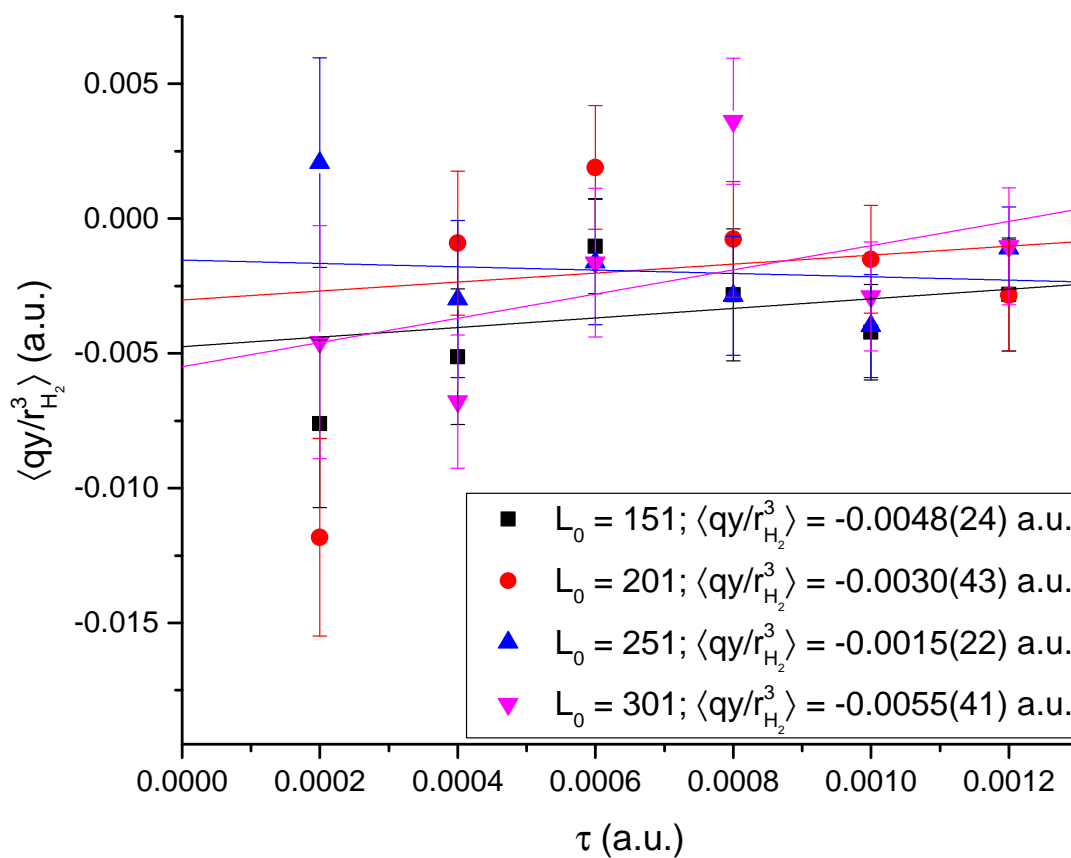


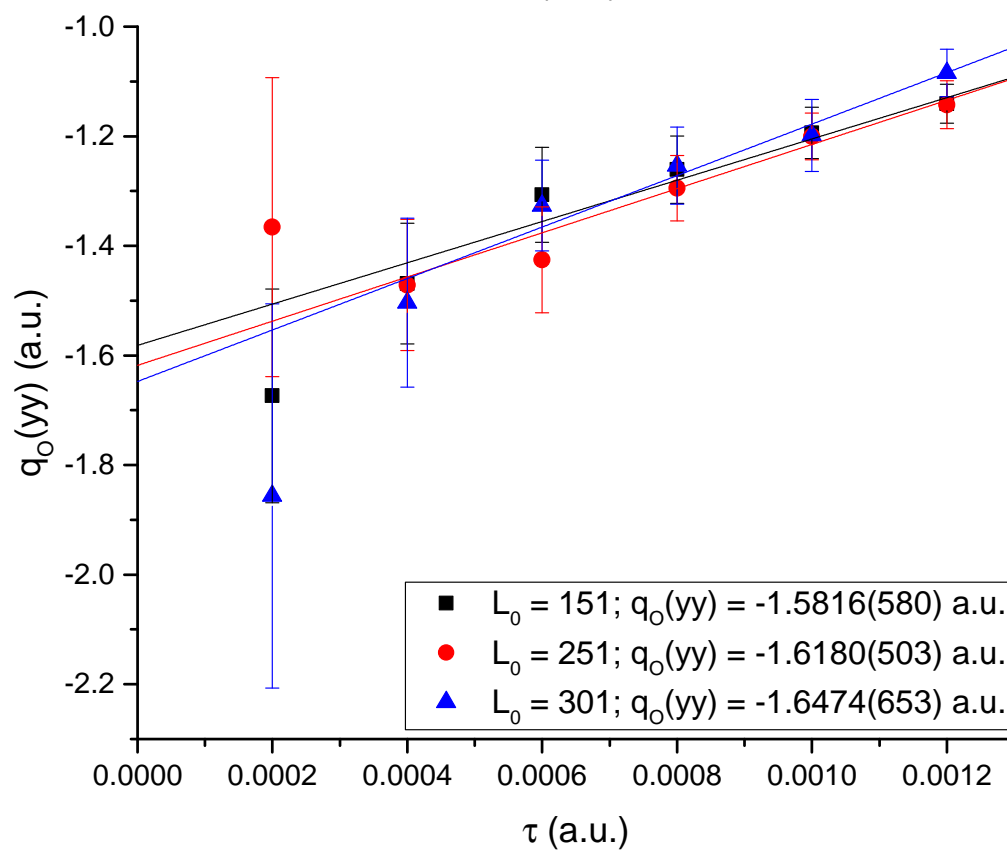
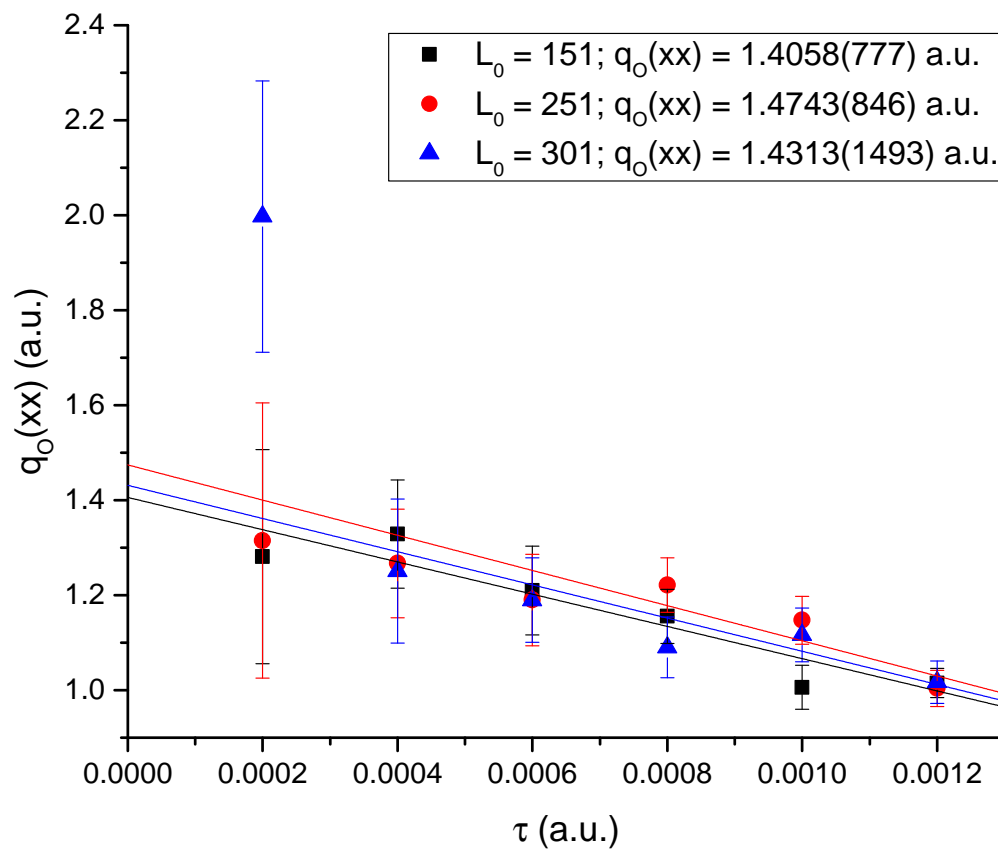


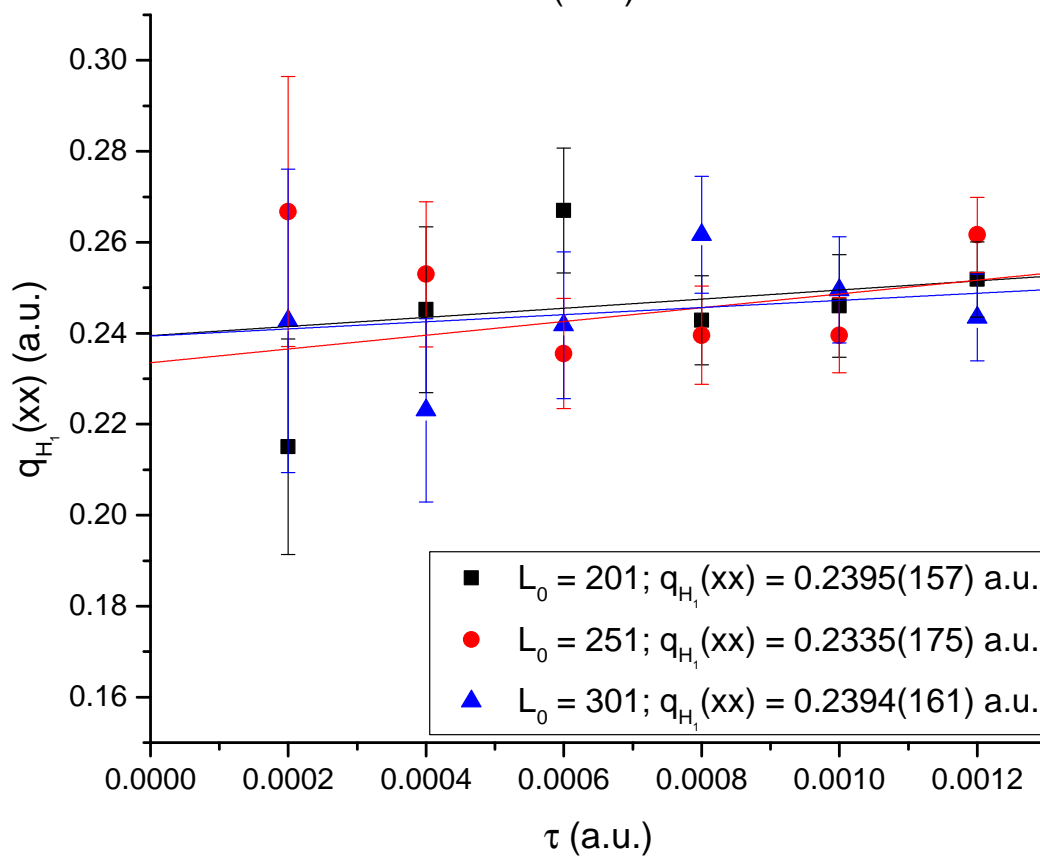
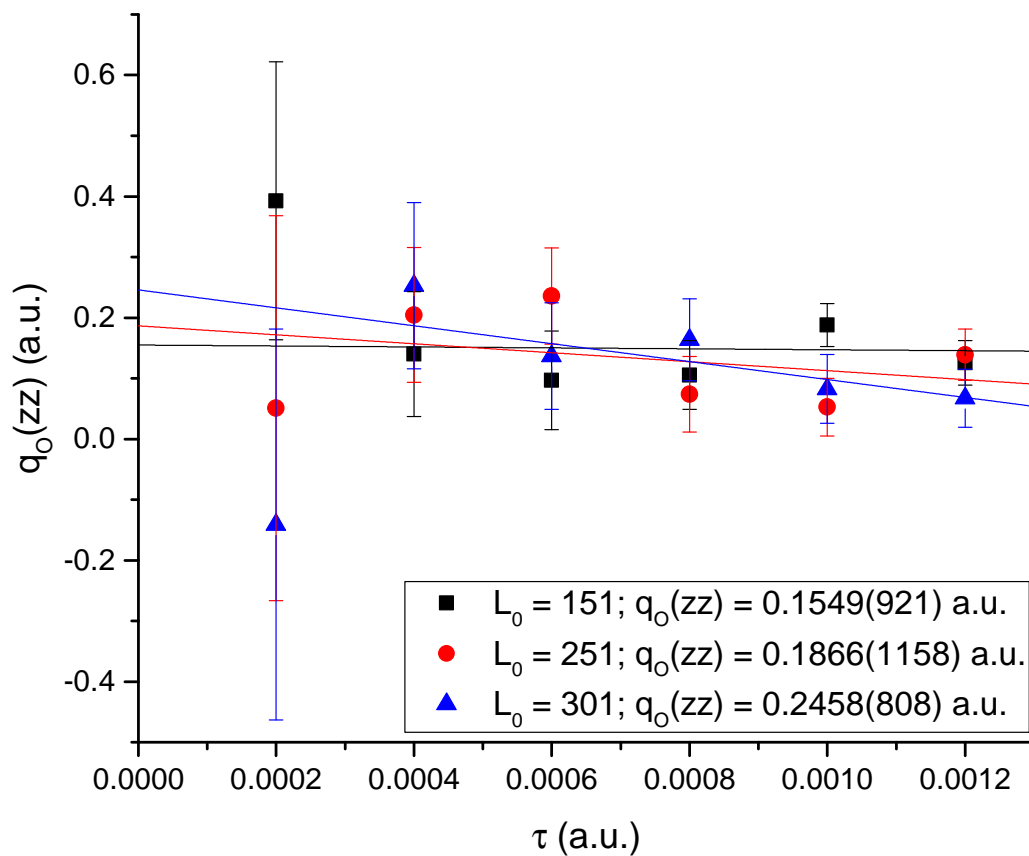


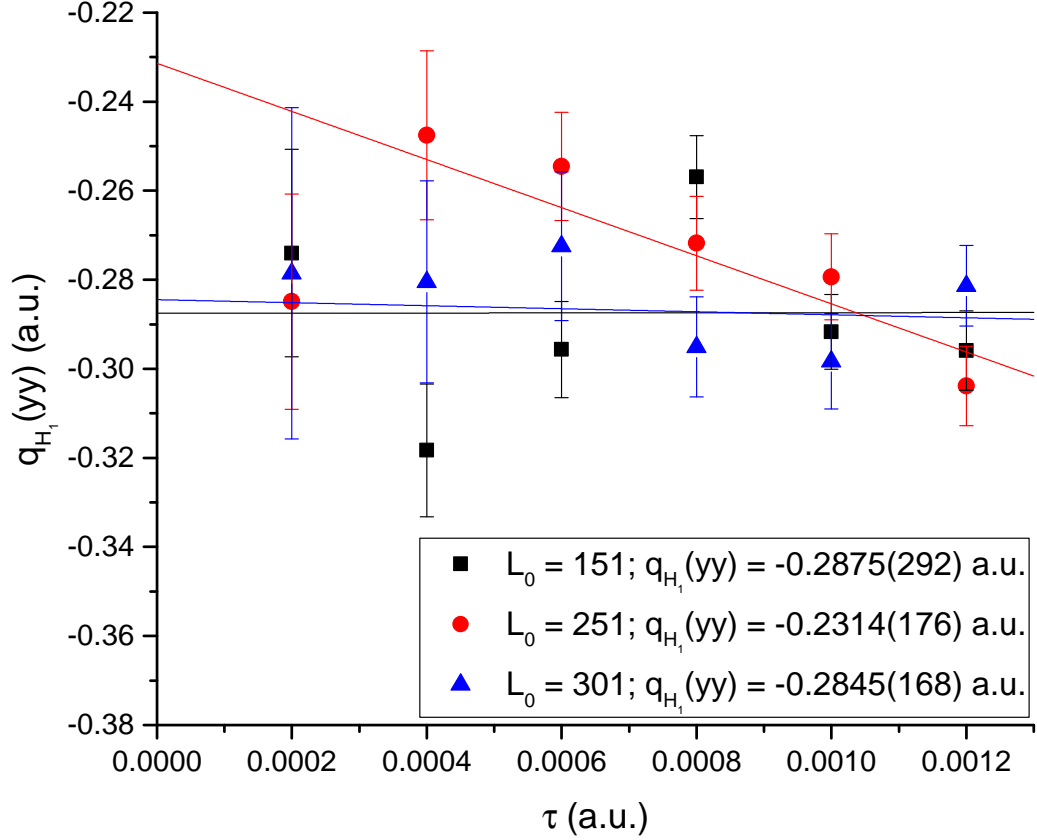
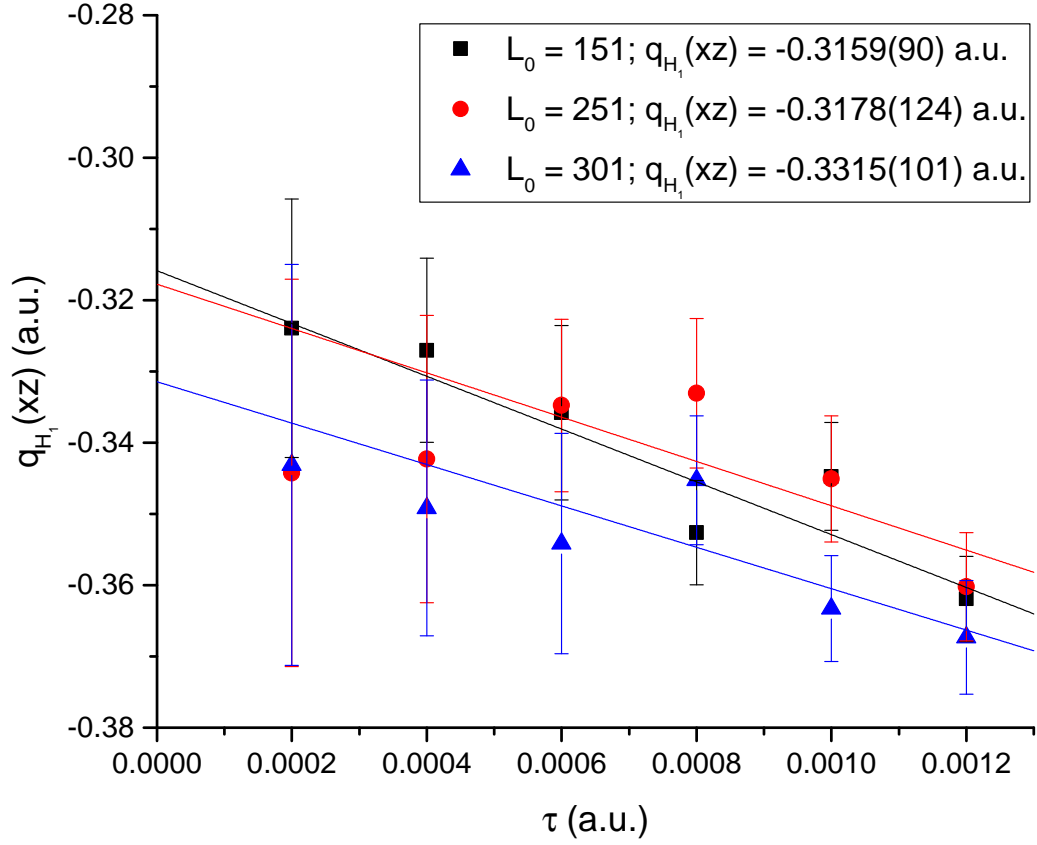


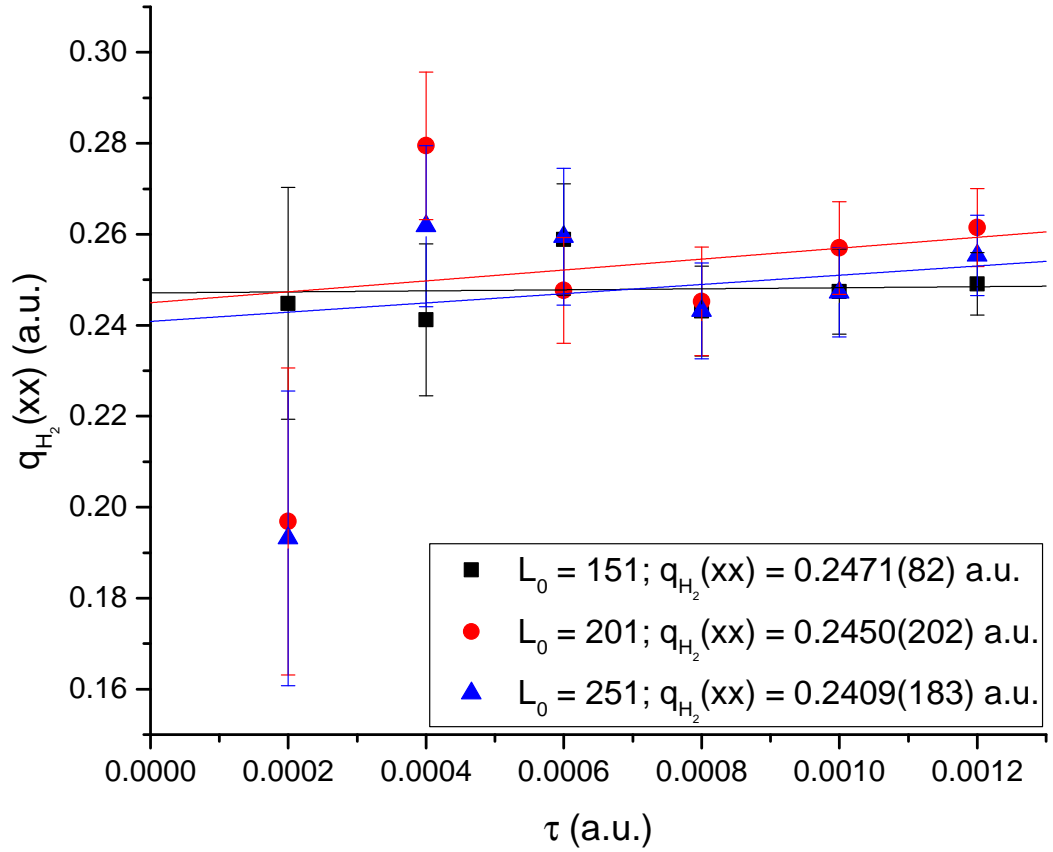
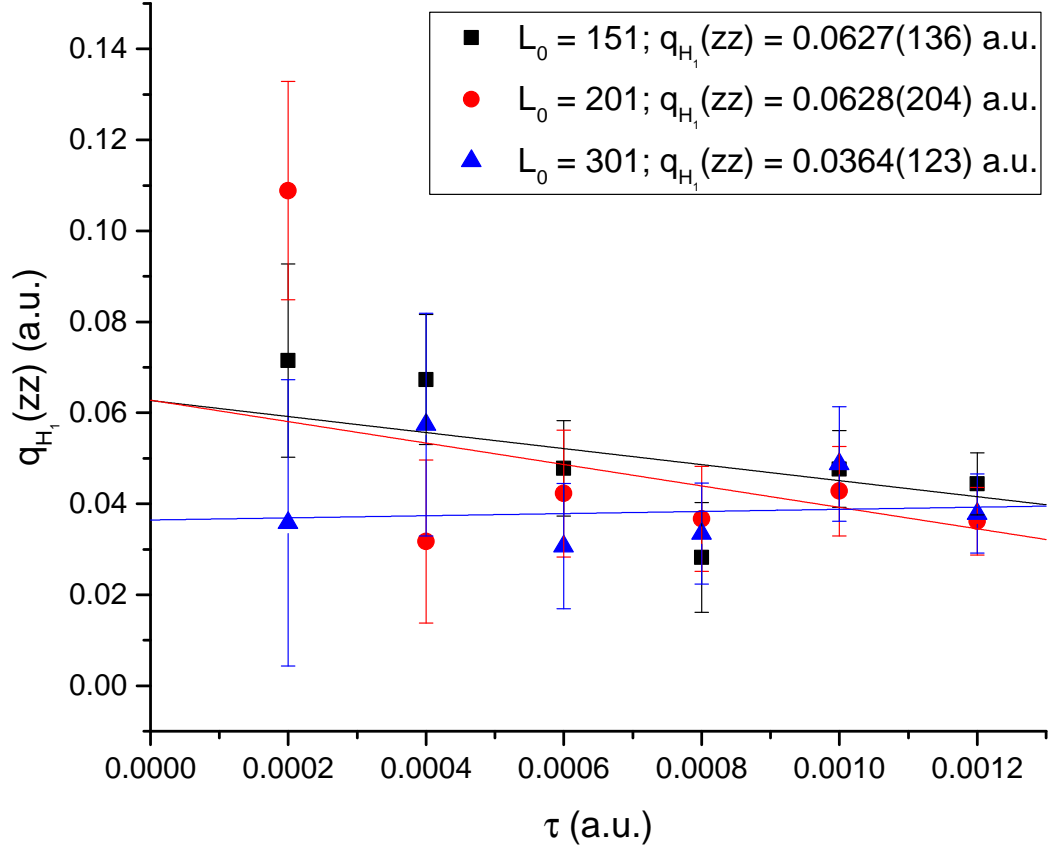


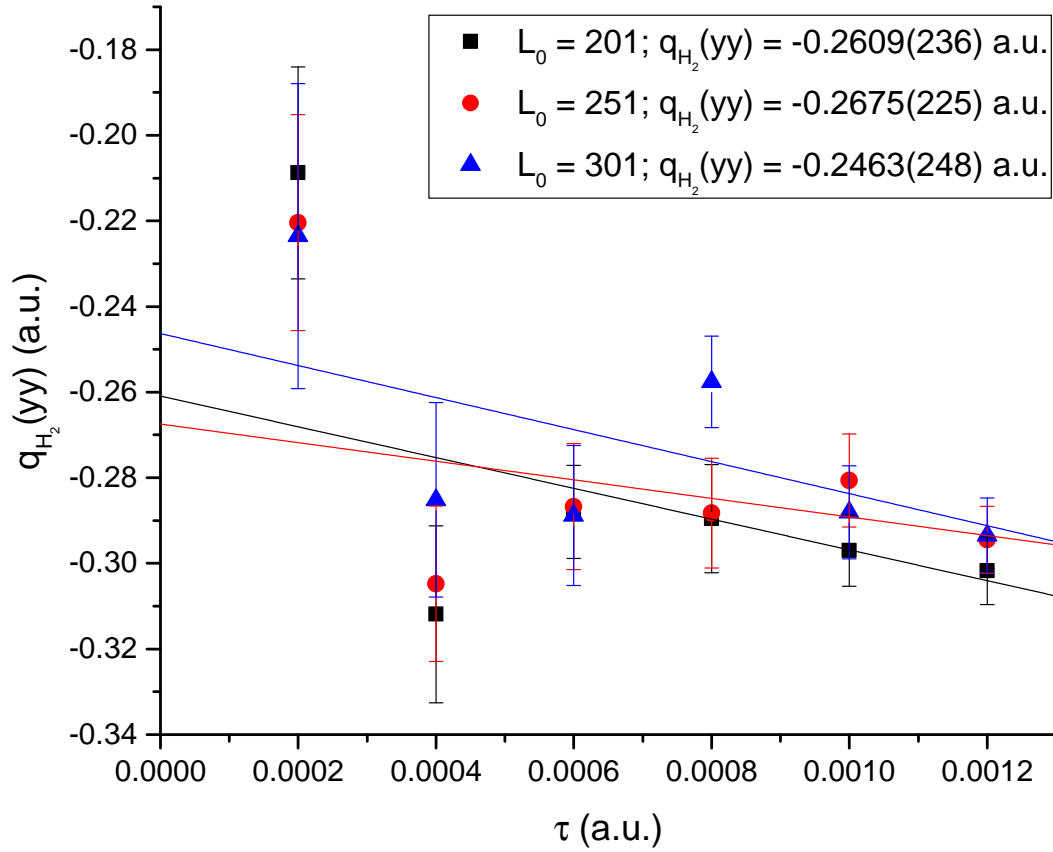
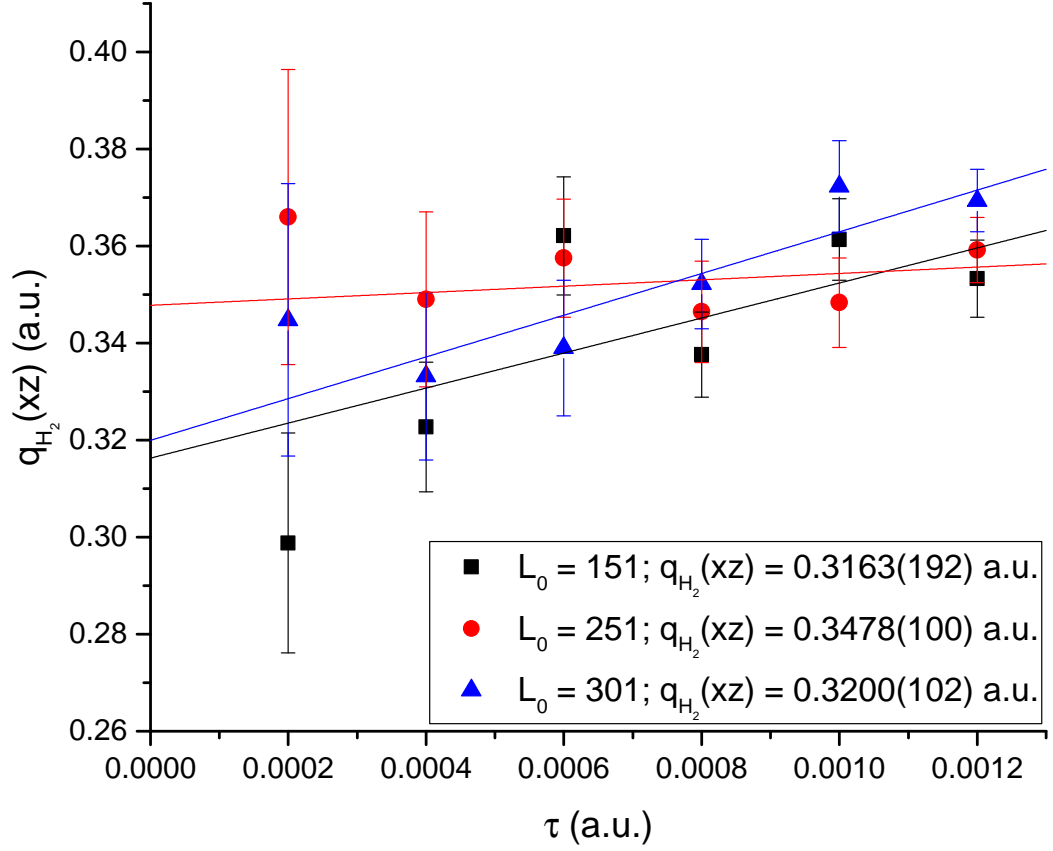




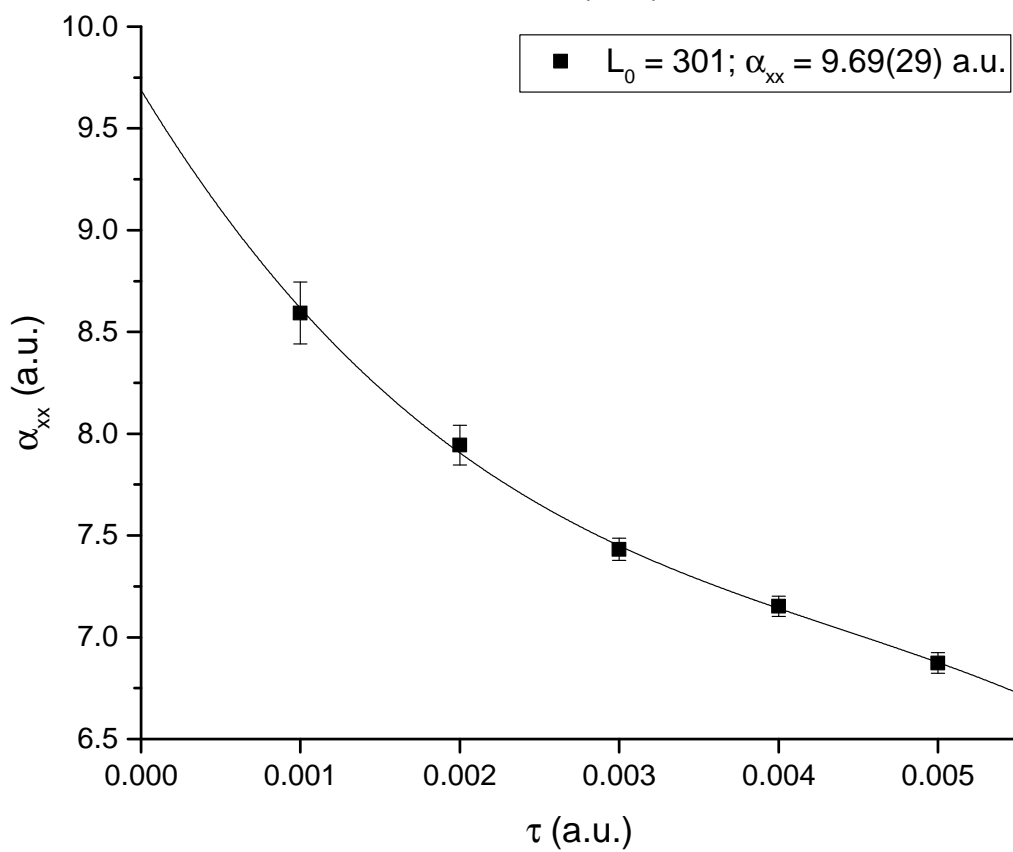
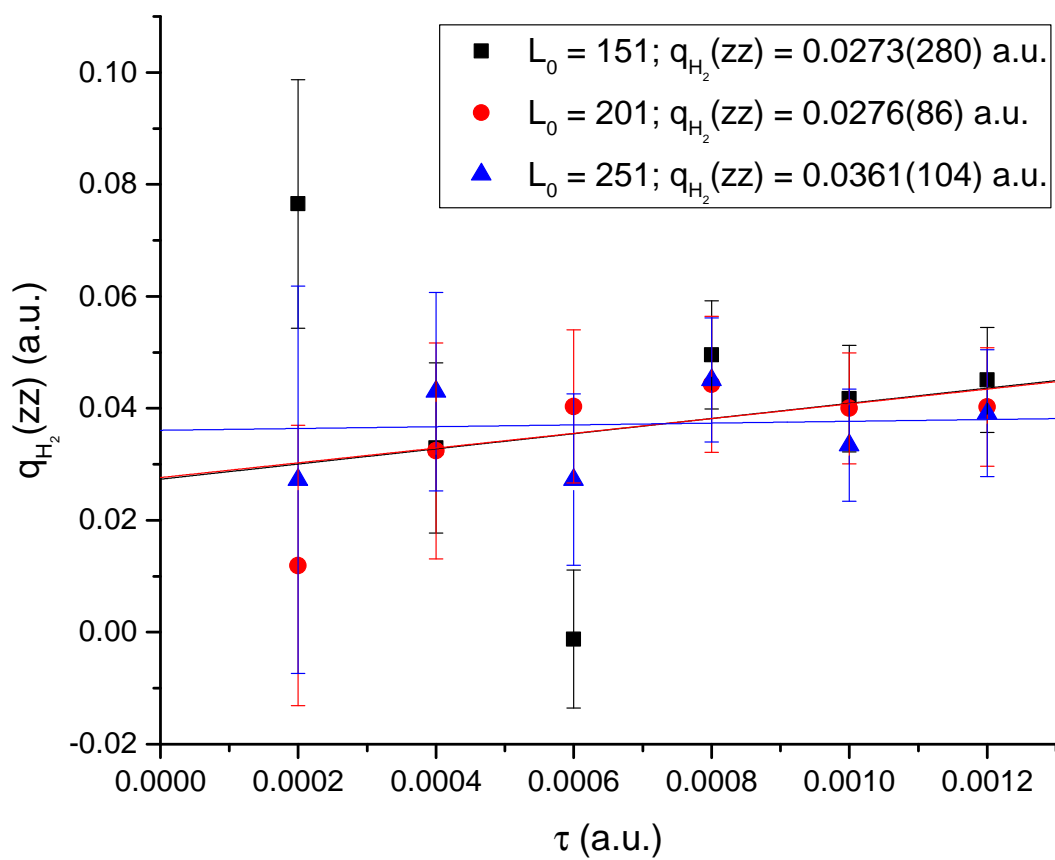


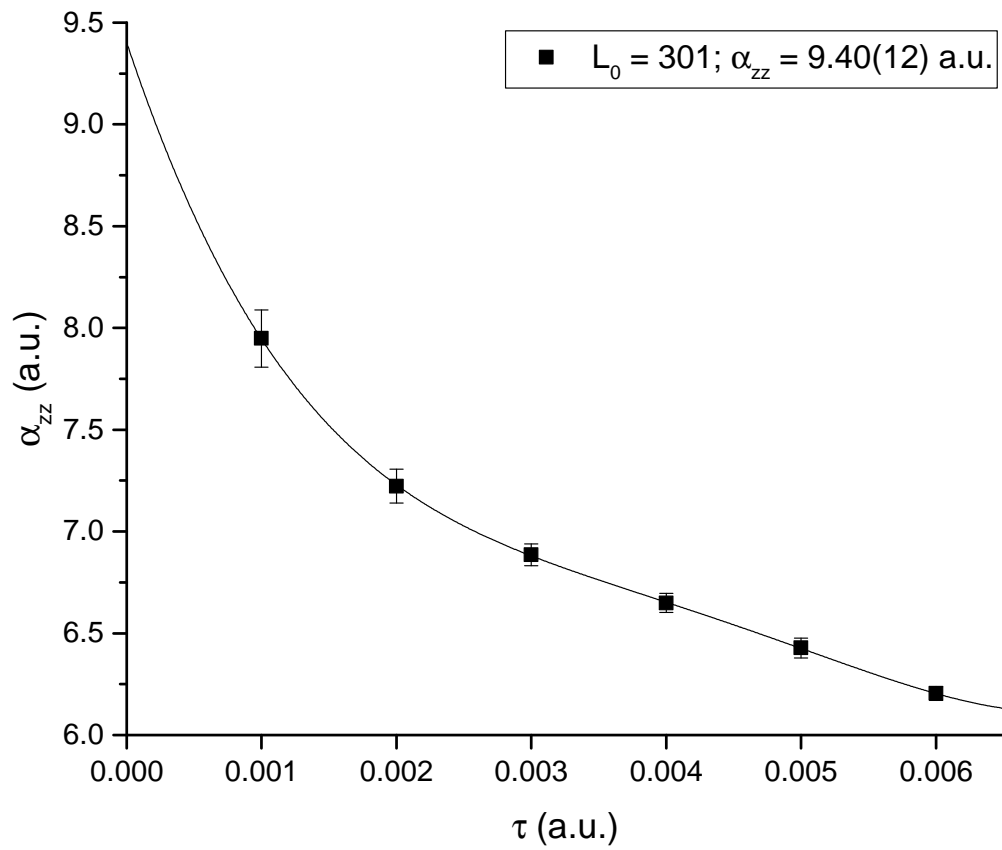
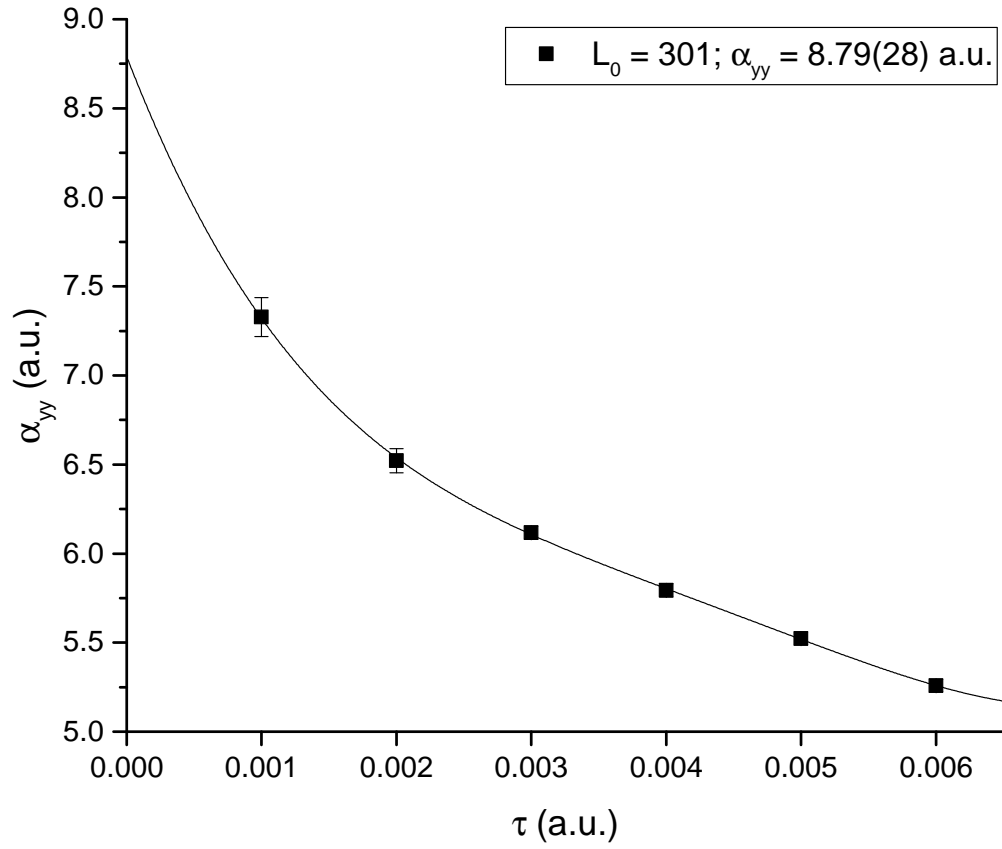












### A.5 Graphs of the Ground-state Energy and Electronic Properties of Ethylene

